ACADEMY OF SCIENCES, U. S. S. R.

УСПЕХИ ФИЗИЧЕСКИХ НАУК

(Advances in the Physical Sciences)

Volume 66

(September—December, 1958)

Translated by the AMERICAN INSTITUTE OF PHYSICS INCORPORATED

NEW YORK, 1959

Editorial

USPEKHI FIZICHESKIKH NAUK is similar to Reviews of Modern Physics in that it is devoted essentially to detailed and exhaustive surveys of various physics subjects. The frequent references made to these survey articles in the Russian and European physics literature are sufficient evidence of their lasting interest. The American Institute of Physics is therefore pleased to add this important periodical to the growing list of its translation journals.

Unlike the Reviews of Modern Physics, "Uspekhi" contains also selected translations of important papers written outside the Soviet Union, reports on scientific conferences, book reviews, personalia, and brief communications on modern ideas or equipment. Each issue of Soviet Physics-USPEKHI will contain all the material contained in *two* issues of its Russian counterpart. Only articles originally written in Russian will be translated.

Needless to say, criticisms and suggestions will be welcomed by the editor.

SOVIET PHYSICS

USPEKHI

A Translation of Uspekhi Fizicheskikh Nauk

SOVIET PHYSICS USPEKHI VOL. 66 (1), NO. 1, pp 1-177 SEPTEMBER-OCTOBER 1958

*RAMAN SPECTRA AND STRUCTURE OF CERTAIN INORGANIC TYPES OF GLASS**

Ya. S. BOBOVICH AND T. P. TULUB

Usp. Fiz. Nauk **66**, 3-41 (September, 1958)

CONTENTS

INTRODUCTION

INORGANIC glass, particularly the silicate glass which is extensively used in science, technology, and household products, belongs, like organic polymers, to the extensive class of amorphous substances. Consequently much similarity should be found in the processes by which both types of substance are formed. Their structures, to some extent, should also be similar.

It is universally known that amorphous substances have no crystal structure. The very fact that some substances cool in an amorphous, glass-like form while others produce crystals can now be explained, in broad outlines, by the classical theory of Tamman.[1]

For a more detailed description of the properties of amorphous substances it is necessary to resort to their kinetic and relaxation characteristics. The physical principles of the corresponding concepts is known to have been formulated by P. P. Kobeko; they have found exhaustive experimental confirmation in the works of P. P. Kobeko and his students,[2] and a mathematical treatment in the works by M. V. Vol'kenshtein and O. B. Ptitsyn.[3]

The microscopic quantity that characterizes the process of vitrification is the viscosity of the melt. At certain temperatures called the vitrification temperatures and denoted by T_g the viscosity becomes exceedingly high ($\sim 10^{13}$ poises), and it becomes increasingly difficult for the molecules to change their orientation. The molecules "do not get a chance" to regroup and an equilibrium state that corresponds to a higher temperature "freezes".

From the molecular point of view, the sharp increase in the viscosity with decreasing temperature of the melt indicates merely an increase in the influence of the intermolecular forces. This increase is due to the reduced mean thermal energy of the molecules. When the ratio of thermal energy to the energy of molecular interaction reaches a

*This article is aimed at a more or less complete survey of work on Raman scattering in silicate glass. Investigations devoted to a study of glass by other physical and physico-chemical methods will be touched upon only as required by the principal purpose of this survey.

certain value, the melt cools in the form of glass.

It is quite obvious that the kinetics of the vitrification process should depend to an exceedingly high degree on the molecular structure of the substance. The more complicated the molecules, i.e., the longer the molecular chains and the more branches they have, the sharper the increase in the viscosity of the melt upon cooling and consequently, the easier the vitrification. Experience shows that this is indeed so. For example, an investigation of the homological series of fatty alcohols has shown that in general the higher terms of the series[2] have the greatest tendency towards vitrification. The inorganic types of glass are no exception. It is known for example that the easiest to cool in the form of glass are silicates with a large deficit of oxygen atoms, corresponding to a content of large and irregular groups of silicon and oxygen atoms in the melt. As the relative number of oxygen atoms increases, the tendency to vitrification decreases. In the limit, when the Si:O ratio becomes exactly 1:4 (orthosilicates), glass is formed only in rare cases. We shall consider this problem in detail later on.

Although the processes of formation of organic polymers and inorganic glass do not differ in principle, nevertheless they are not identical in detail. A very important specific feature of certain types of inorganic glass, particularly complex silicates, is that they are formed of metallic and non-metallic oxides, which give rise to a great variation in the character and strengths of the bonds. This leads to a great variety in the possible structure of the resultant amorphous body, a structure essentially determined both by purely geometric factors (the radii of the ions) and by their charges, by the degree of covalence (ionization), and by the binding energy.

Scientists who investigate the basic features of the structure of inorganic glass all have a unilateral approach to the problem—a tendency to emphasize some one of the aforementioned specific features of the given systems, rather than attempt to examine the problem as a whole, from a general physical review of the nature of the chemical bond and the structure of amorphous bodies.

The most frequent approach to the problem of structure of inorganic glass, from the point of view of crystal chemistry of ionic crystals, is based on the analogy between the structure of the crystal and the corresponding glass.

The simplest premise on which such an examination is based is that only electrostatic forces act between the ions. Actually, however, one must not lose sight of other forces, for example those due to the partially covalent character of the bond. Under certain cases the bond must be predominantly covalent. Then the purely-geometric factors no longer play a decisive role. We shall return to this problem later on.

Goldschmidt[4] already tried to relate the fact that certain oxides such as SiO_2 or B_2O_3 readily produce glass while others, say Al_2O_3 cannot be obtained in glass form under any circumstance to the ratio of the radii $R_A R_O$ in the oxide A_nO_n. It was found that this ratio ranges from 0.2 to 0.4 in all oxides capable of vitrification.

Cases are known, however, when this rule is not obeyed. Thus, the value of R_{Be}/R_O for BeO is 0.3, i.e., it lies in the required range, but nevertheless no glass-like BeO has yet been produced.

A similar criterion was applied by Sun and Silverman to multi-component glass.[5] Pincus attempted to characterize the tendency to glass formation by means of another quantity, the ionic potential, defined as a ratio of the ion charge to its radius.[6] Here two physical parameters are already taken into consideration, the ionic charge and its radius. Thus Pincus' attempt was a certain step forward.

Many substantial investigations of the glassy state, also based on crystal-chemical analogies, have been made by Dietzel.[7, 8] He has shown that certain properties of glass (viscosity, ability to flake or crystallize, etc.) can be described with the aid of a quantity called the "field intensity" of the cations (K^+, Na^+, Ca^{++}, Pb^{++}, Ba^{++}, etc.), specified at the distance at which the anion of the oxygen of the glass-forming oxide is located (and not at the distance R of the cation, as proposed by several earlier workers), and characterizing the Coulomb interaction between the cation and a single charge of the oxygen. All the elements of interest in glass melting were arranged by Dietzel in a series according to the value of the "field intensity." The position of the element in this series was compared with its tendency to become incorporated in the silicon-oxygen lattice of the glass.* It turns out that the stronger the field of the cation, the less pronounced is this tendency. Very "strong" cations are more prone to formation of their own lattice with the oxygen atoms. The corresponding types of glass flake readily. On the contrary, glass containing "weak" cations in coordination with the oxygens of the SiO_4 tetrahedra are homogeneous and do not flake.

It has already been noted above that many prop-

*Dietzel does not distinguish between ionic or covalent bonds of elements and the structural lattice of the glass.

TABLE I. Differences in electronegative potential and in the degree of ionization of the A-O bonds

Serial number	A–O bond	Difference in electronegative potential, $X_A - X_O$	Degree of ionization in percent	Serial number	A–O bond	Difference in electronegative potential, $X_A - X_O$	Degree of ionization in percent
1	N—O	0.5	6	12	Be—O	2.0	63
2	C—O	1.0	22	13	Mg—O	2.3	73
3	S—O	1.0	22	14	Ca—O	2.5	79
4	B—O	1.5	44	15	Sr—O	2.5	79
5	As—O	1.5	44	16	Ba—O	2.6	81
6	Si—O	1.7	50	17	Li—O	2.5	79
7	Ge—O	1.7	50	18	Na—O	2.6	81
8	Sb—O	1.7	50	19	K—O	2.7	83
9	Ti—O	1.9	58	20	Rb—O	2.7	83
10	Zr—O	1.9	58	21	Cs—O	2.8	85
11	Al—O	2.0	63				

erties of glass are qualitatively explained within the framework of Dietzel's ideas. Yet, an explanation of many factors involves difficulties. Thus, for example, the properties of glass containing such cations as Pb^{++}, Cd^{++}, Zn^{++}, and certain others (with electron shells unlike those of the inert gases), are incompatible with the positions of these cations in the table of the cited papers. Dietzel does not give any satisfactory explanation for these deviations.

Summarizing this brief statement of Dietzel's ideas, we must above all emphasize the undoubted success that he attained in describing various properties of glass. This success is evidence of the substantial role of electrostatic forces in the vitrification process. It would be erroneous, however, to expect complete agreement between these ideas and experiment. In fact, the forces acting between the individual structural elements of the glass cannot be ascribed exclusively to electrostatic interactions. One must always take into account the partially covalent character of the bonds.

To understand more clearly the relation between these two factors we tabulate (Table 1), on the basis of literature data on the value of the electronegative potential,[9] the values of the degree of ionization of various A-O bonds. It is seen from the data of the table that in most of the cases cited the degree of ionization is less than 50% or slightly above it. The bonds that have an ionic character are for the most part essentially those formed with participation of the alkali and alkali-earth elements. The first three elements of the table form with oxygen a predominantly covalent bond. The calculation of a "field intensity" for such bonds, as was done in one of Dietzel's papers,[8] therefore hardly has any physical meaning.

The fact that many properties, nevertheless, are in satisfactory agreement with Dietzel's conclusions can be explained apparently by the following considerations. A comparison of the degree of ionization of the bonds with the field intensity of the cations, arranged in the same order, shows that they vary in opposite directions, the variation in field intensity being more prominent. For example, for potassium these quantities are 83% and 0.13, for boron they are 44% and (depending on the coordination number) 1.34 to 1.63 respectively.* We thus see an approximately twofold reduction in the degree of ionization of the bond, but an increase in the field strength by one order of magnitude. In other cases we observe qualitatively the same pattern. Consequently, the relative role of the electrostatic interaction in the formation of the A-O bond exhibits a net increase, although a slower one than that of the field intensity in Dietzel's table of elements.

It is hardly worth while to compare directly the vitrification ability with the degree of covalence of the bonds, although some grounds for this are usually seen in the fact that the directional character of most covalent bonds contributes to the formation of a lattice or in any case, to large and irregular groups. For many oxides this is apparently indeed so (SiO_2, B_2O_3, GeO_2, etc.) There are, however, exceptions to this rule. A striking example of this kind is BeF_2. According to Pauling's data on electronegative potentials, the Be-F bonds have a ionization degree of approximately 80%.[9] Nevertheless it is relatively easy to obtain BeF_2 in the form of glass.[10] Consequently, the covalent character of the bonds does not always play a substantial role in the vitrification process. Any criteria based on calculation of the covalence, although usable, must therefore be approached with care.[11]

Deserving of attention is the fact that the litera-

*The field intensity will be given here and throughout in units of the elementary charge.

ture contains contradictory opinions concerning the character of the Be-F bonds. An estimate made on the usual assumptions, as noted above, shows that in this case the ionic character of the bond predominates considerably. This is a natural result, since the F atom is the most electronegative of all elements. Yet, according to Neumann and Richter,[12] a melt of BeF_2 is characterized by very small electronegative potential, which, so to speak, demonstrates the substantially non-ionic character of the Be-F bonds. In the paper cited above[10] Hagg calls attention to the acid properties of BeF_2, in contrast to the base properties of BeO. This circumstance would appear to confirm the results of Neumann and Richter.[12]

One might assume that an important role in vitrification processes is played not by the type of the bond between the metal and the oxygen, but rather by the strength of this bond. Actually, it was noted earlier that vitrification is caused by a sharp increase in the viscosity near the temperature T_g, a viscosity which in turn is directly dependent on the magnitude and irregular character of the molecular groups. In the final analysis it is to a certain extent immaterial what forces cause the formation of such groups. It is merely important that "seams" can appear between the individual molecules and that these "seams" are sufficiently strong.

To verify the extent to which these concepts are correct in the case of BeF_2, we calculated the Coulomb attraction between the Be^{++} and F^- ions and compared the data with those given by Dietzel for the A-O bonds. It turned out that the Coulomb attraction is in this case approximately 0.7 (in arbitrary units). For glass-forming oxides (P_2O_5 SiO_2 and B_2O_3) the corresponding values, using the same units are according to Dietzel[8] between approximately 2 and 1.35, i.e., substantially larger.* It is thus impossible to explain the anomalous behavior of BeF_2 on the basis of the primitive ideas concerning the strength of the Be-F ionic bonds.

In a recently published, exceedingly interesting work by Winter-Klein[13] the ability of vitrification is compared with the structure of the external electron shells of the atoms. The author shows by means of numerous examples that a quantitative criterion of this kind can be the number of external p electrons per single atom. The greater this number, the stronger the vitrification ability. If this number is less than or equal to two, no vitrification takes place at all.

It is most natural to attempt to express the rules established in the work by Winter-Klein in terms of the role of the strength and directional character of the covalent bonds.

In fact, it is well known from quantum-mechanical theory, that covalent bonds between two atoms are stronger, the more their orbits overlap. The p orbits overlap the s orbits more strongly than the s orbits overlap among themselves. The p-p orbit overlap is even stronger. If the strength of the s-s bond is taken to be unity, then the strength of the s-p bond becomes 1.73, and that of the p-p bond 3. The bonds become even stronger because of the interpenetration of the orbits. Differing greatly from the s orbits in their angular distribution, the p orbits determine furthermore the orientation of the bonds.[9]

Thus, generalizing the observations of Winter-Klein, it would appear that one can state that the number of p electrons per single atom of a given molecule characterizes the "average" strength and directional character of the bond. It is, indeed, these parameters of the bonds that can become important in vitrification processes.

If this is so, then one cannot lose sight of the fact that other outer orbits, such as the d and f orbits, can play an analogous role in the strengthening of the bonds and in making the bonds directional. By way of examples in which the d and f orbits must be considered, we indicate such important elements in glass production as titanium and vanadium, and also zirconium and molybdenum, for which the outer orbits are 3d and 4f respectively.

Under certain conditions these orbits, owing to interpenetration, may be even more effective than the p orbits.

Let us attempt to apply the above considerations to certain simple systems.

From the data of Table II it is seen that the application of the Winter-Klein criterion to certain simple systems is not in satisfactory agreement with experiment. Thus, for example, the boric anhydride is characterized by a criterion value of 2.8 while quartz has a value 3.3. Yet the former apparently produces glass more readily, as evidenced by the fact that it does not crystallize under almost any conditions; V_2O_5, which has the highest value of this criterion, produces no glass at all.

Thus, a criterion based on taking into account only the structure of the external electron shells of the interacting atoms cannot be considered universal.

*Here, as in further calculations of this kind, small corrections for the effective coordination number are generally disregarded.

TABLE II. Estimate of certain simple systems with the aid of the Winter-Klein vitrification criterion[13]

Serial number	System	Number of atoms	Number of p and d electrons	Number of p and d electrons per atom	Remarks
1	B_2O_3	5	14	2.8	Vitrifies readily Practically does not crystallize
2	SiO_2	3	10	3.3	Vitrifies readily Crystallizes
3	TiO_2	3	10	3.3	Does not vitrify
4	Bi_2O_3	5	18	3.6	Does not vitrify
5	V_2O_5	7	26	3.7	Does not vitrify

A special position is occupied in the rather extensive literature on the structure of glass by the principles postulated by Zachariasen.[14] The starting point in these principles is the assumption that the energy of the glass lattice is commensurate with the energy of the corresponding crystal. More accurately expressed, it exceeds it slightly, which explains, generally speaking, the merely weak tendency of glass to pass through crystallization. According to Zachariasen, a non-periodic lattice satisfies this condition.

For a long time, the view of Zachariasen that glass is a non-periodic lattice was contrasted with the crystallite concepts which, as is well known, were first formulated by A. A. Lebedev[15] and later on intensely developed by his students and followers.[16]

According to the initial version of the crystallite theory glass was considered, from the structural point of view, as a mass of very small crystalline formations of various silicates, whose nature is directly related to the composition of the glass. Subsequently the opinions of most investigators leaned towards the idea that the concept of crystallites should not be reduced to that of microcrystals with phase boundaries, but the crystallite is more correctly taken to be a regularized nearest-order region inside a general irregular and non-periodic lattice.[16] The quantitative relation between the regular and irregular structure of glass is still unclear.

Thus, the contrasting of these points of view has lost its meaning. Either one can be considered correct, each showing different aspects of the general problem.

In connection with this we note that the ability of crystallite formation is not an exclusive property of silicate or inorganic glass. Organic polymers also display such a property. There, too, the crystallites are assumed to be nearest-order regions. Nevertheless, the presence of such regions did not in any way affect the conventional theoretical treatment of the structure of organic polymers.[2]

Summarizing briefly the papers surveyed above, which by far do not exhaust the literature on this probelm, we can state that not one of the criteria considered, which in the opinion of their authors characterize the ability of various elements to produce simple and complex types of glass, is in satisfactory agreement with experiment. This is why direct methods of investigating the structure of various glass-like bodies, which would permit the establishing of an atomic-molecular "architecture" of glass and which would also permit estimating the character of the bond between its individual structural elements, become particularly important. Spectroscopic methods can play a substantial role in this respect.

1. FEATURES OF THE RAMAN LIGHT-SCATTERING METHOD AS APPLIED TO THE INVESTIGATION OF GLASS

All the necessary information on vibrational spectra are contained in two monographs, references 17 and 18. We shall not discuss this problem.

There is also a rather extensive literature[19] devoted to experimental means and methods used to obtain and investigate Raman spectra. Not taking it upon ourselves to explain all aspects of this great and specialized problem, we confine ourselves merely to a few remarks connected with the specific nature of scattering in glass.

The considerable difficulties in obtaining scattering spectra in glass are due to the low intensity of the bands. One of the methods of reinforcing the bands consists of intensifying the illumination of the specimen. However, the use of several

high-power high-pressure mercury lamps for this purpose, the usual procedure in the investigation of liquids, hardly settles this problem, for the spectrum of these lamps contains a rather intense continuous background in addition to sharp lines. This background always interferes with the observation of weak scattering bands. It is particularly harmful in the study of glass, since the specimens as a rule have defects (knots, bubbles, etc.) which reinforce the parasitic scattering.

Most effective in the investigation of glass are high-power, low-pressure mercury lamps. They were first used for this purpose in the United States[20] and then by the authors of reference 21. These lamps are not only practically free of a continuous background, but at the same time produce a very strong radiation flux. A description of low-pressure mercury lamps can be found in the original papers.[22]

Light can be resolved and recorded photographically either with prism spectrographs or with diffraction-grating spectrographs. When glass is investigated, it is essential that the spectral instruments have a sufficiently high aperture. Otherwise the exposure time needed to photograph the spectra may reach many hours, and it becomes difficult not only to measure but even to estimate the intensities and polarizations of the bands.

The experimental conditions are considerably facilitated if photography is replaced by the photoelectric method of spectrum recording. In this case, firstly, the time required to obtain the results is considerably shortened and, secondly, exact quantitative measurements become possible.[21, 23, 24]

Unlike the photographic method, the photoelectric method is based on the use of a different spectral instrument, the monochromator. In principle the monochromator is essentially a spectrograph in which the photographic camera is replaced by an exit slit and its dispersing element (prism or diffraction grating) rotates. Thanks to this, the exit slit separates, in sequence, the radiation of a given spectral composition. After separation, the radiation is made to fall on a photoelectronic multiplier. The electric signal is amplified with a vacuum tube amplifier and is usually fed to a vacuum-tube recording potentiometer, which displays the spectrum on a paper chart.

The considerable noise found in the investigation of low-grade glass specimens is due to the fact that the defects in the specimens intensify greatly the mercury lines in the scattered light. These lines, upon entering the spectral instrument, produce an additional parasitic background. To eliminate this noise the authors of references 25 to 27 used the principle of double monochromatization of light. Double monochromatization of light, combined with the use of a low pressure lamp, has reduced the continuous background in such a radical manner, that it becomes possible to investigate, even in very low-grade specimens, the low frequency region of the spectrum directly adjacent to the exciting line.

Let us now touch briefly on the problem of frequency measurement and the degree of deplorization of the bands in the spectra. In the photographic method of measurement the frequencies are obtained by comparing the photographs of the spectra. In this case, generally speaking, great accuracy can be obtained. However, the bands in the Raman spectra of glass are broad and diffused. When measuring the frequencies of such bands difficulties usually arise, owing to a certain degree of arbitrariness in the aiming of the cross-hair of the comparator on the intensity maximum of the band. This is why frequency data obtained photographically are, as a rule, inaccurate. Yet the photoelectric technique of recording the spectra, being free of this shortcoming, is connected with a different source of possible errors in the measurement of frequencies, namely the non-linearity in the operation of the chart-advancing mechanism of the recording potentiometer. To exclude this source of errors, the authors of reference 25 to 27 have decided on a compromise: they measured the frequency visually, following the motion of the recording-potentiometer pen and fixing its position at the maximum of the band by means of a revolution counter, which was part of the double monochromator employed.[28] The greatest measurement error did not exceed 4 cm^{-1}, which is quite acceptable when studying scattering spectra in glass.

As mentioned above, the photographic polarization measurements encounter great experimental difficulties. These difficulties are due to the errors in the photometry of the emulsions. Since the spectra of the glass are of low intensity, the photometry is carried out under very unfavorable conditions (large-grain emulsions) and the sought quantities (the degree of depolarization of the bands) vary only over a narrow range (from 0 to 6 or 7), the accuracy of such measurements cannot be satisfactory. No wonder, then, that the literature contains only two known papers in which an estimate is made by photographic means (and a qualitative one to boot) of the state of polarization of bands in scattering spectra of certain glass-like bodies.

Measurement of the degrees of depolarization becomes further complicated by the specific re-

quirement that the radiation illuminating the object must have a definite direction. This requirement can be satisfied only if the aperture of the apparatus is reduced below normal. If illuminating devices with large apertures are used in the "classical" procedure of measuring the degree of depolarization, in which both components are obtained simultaneously (for example, with the aid of a Wollaston prism in the beam of scattered light), the increase in the aperture ratio is achieved at the expense of reducing the already narrow range of possible values of the degree of depolarization of the apparatus.

It is much better to carry out polarization measurements using a procedure based on the use of polaroid films of different orientations in the illuminating beam, so that the specimen can be illuminated with wide beams. This procedure is particularly convenient if photoelectric recording of the spectra is used.[29] This procedure also permits accurate intensity measurements. Without going into details, we merely note that the exact values of the degree of depolarization of the beams is determined in this procedure with the aid of a calibration curve, which can be readily plotted beforehand. In those cases where exact quantitative data are not required (for example when solving most structural problems), it is possible to dispense with the calibration curve. This is exactly what the authors of references 21 and 24 to 27 have done.

Polarization methods in such objects as glass involve still further difficulties owing to the presence of strains of various intensities in the specimen, which are capable of distorting the polarization state of the bands (particularly of the strongly-polarized ones). Strains are expected to exert the maximum influence in readily-crystallized specimens, the production of which in the glassy state is possible only through abrupt cooling. This is why all quantitative comparisons of different specimens are doubtful or in any case should be carried out with care. The authors of references 21 and 24 to 27 restricted themselves only to qualitative estimates.

2. RESULTS OF EXPERIMENTAL STUDY OF RAMAN SCATTERING IN GLASS

A. General Character of Scattering, Multi-Component Silicate Glass and Fused Quartz

Most of the earlier investigators of Raman scattering in glass concerned themselves with complicated, multi-component systems. Nevertheless, they succeeded in establishing the specific features of this phenomenon. Therefore, adhering where possible to the chronological order, we begin the survey with these investigations.

The first attempts to investigate Raman spectra in silicate glass date back to 1928.[30-33] However, the results then obtained were contradictory and gave no definite indication even whether Raman scattering spectra exist at all in glass, let alone the character of this scattering. Thus, for example, Raman[31] and Raman and Krishman[32] suggest that Raman spectra in glass are characterized by diffuse bands, while Pringsheim and Rosen,[33] in spite of very prolonged exposures, could not detect even traces of bands in the investigation of fused quartz. One can consider the first reliable and clear data to be those obtained in 1929 by E. F. Gross and M. F. Romanova,[34] who investigated specimens of borosilicate crown glass, light and heavy flint glass, and fused and crystalline quartz. They actually observed broad smeared bands in the spectra of these substances, approximately in the same positions as the lines of crystalline quartz, but varying greatly in intensity from substance to substance. The spectrum of crown glass turned out to be much weaker than that of flint glass, and this was confirmed by subsequent investigations.

The character of the spectra (smearing of the bands), together with x-ray structural-analysis data, proves without doubt, in the Gross' and Romanova's opinion,[34] the amorphous structure of the investigated glass. The fact that the frequencies in the spectra of the glass coincide with those in crystalline quartz proves that the region amenable to investigation does not contain the frequencies that are characteristic of the vibrations of the quartz crystal lattice, but contains only the frequencies belonging to the internal vibrations of the individual SiO_2 groups.

In connection with this we note the following. The concept of crystal-lattice vibrations is applicable, strictly speaking, to molecular crystals, for which the internal vibrations can be separated from the external ones. A quartz crystal is essentially a gigantic molecule made up of atoms joined by bonds of equal type; this molecule is characterized by a regular alternation of the SiO_4 tetrahedra. The concept of vibrations of such a crystalline lattice therefore ceases to be precise.

The most complete and systematic research on glass of various complicated compositions was carried out by Bhagavantam,[35] Kujumzelis,[36, 37] Langenberg,[38] and Norris.[39] Let us dwell on some of these investigations.

Kujumzelis[36, 37] investigated 19 specimens of multi-component glass, including a system that

contains predominantly PbO and SiO_2 in various ratios, potassium-phosphate glass with a small addition of other cations, glass-like B_2O_3 and SiO_2, and certain other types of glass. In spite of the complexity of composition of most specimens, he succeeded, nevertheless, in establishing many general features of scattering in glass and in propounding certain ideas concerning the structure of glass, some of which make up an acceptable working hypothesis.

Kujumzelis observed for each of the investigated specimens a characteristic spectrum that consists, in agreement with the observations of E. F. Gross and M. F. Romanova,[34] of broad bands of different intensities. Perhaps the most interesting and newest detail that he has found is a continuous Raman spectrum in lead glass, in the form of a broad continuous band with maxima directly adjacent to the exciting line; this band has a sharp boundary at approximately 550 cm^{-1}.

Since the spectra of the glass do not exhibit a clear similarity with the spectrum of fused quartz and depend, furthermore, on the composition of the glass, Kujumzelis proposes that glass contains, in addition to coupled SiO_4 tetrahedra, other groups, groups consisting primarily of atoms (anions) of oxygen, which can be contained inside the cations of the metals and of the glass-forming oxides. Here the author leans on several remarks by Bragg, made in his well-known monograph on the structure of silicates.[40] Two of the bands, with frequencies of 610 and 800 cm^{-1}, which retain their positions almost exactly in the spectra of all types of glass including fused quartz, are ascribed by the author to those vibrations in which only oxygen atoms participate, for example six-term rings made of oxygen atoms.

Kujumzelis approaches the explanation of the aforementioned continuous scattering from the point of view of Zachariasen, and assumes that this scattering is due just to the superposition of the vibrations of the irregular and anomalous lattice of SiO_4 tetrahedra, a lattice which is a characteristic of glass. The sharp boundary of the continuous scattering band should obviously be compared with the vibrations of the individual SiO_4 tetrahedra as a whole.

While the interpretation of the continuous scattering and the assumed existence of different groupings in the glass is sufficiently plausible and can be considered acceptable, the attempt to detail the concept of such groupings appears to us unfounded and speculative.

Langenberg[38] investigated even more glass samples, also with complex but controlled composition. He found certain insubstantial corrections to the data obtained by Kujumzelis.[36,37] In the interpretation of his own results, Langenberg starts with the assumption the spectra of the glass contain frequencies of isolated tetrahedra, the values of which can be found in the paper by Schafer, Matossi and Wirtz.[41] The particular fact that the bands in glass acquire a diffused character (compared with crystalline quartz) is indeed evidence, according to Langenberg, of variations in the symmetry of the tetrahedra and the splitting of the degenerate vibrations of the glass.

One can hardly agree with an interpretation of the spectrum of the silicates and silicate glass based on the frequencies of an isolated SiO_4 tetrahedron. It is impossible, even in rough approximation, to separate the free tetrahedra in any glass (with the possible exception of orthosilicate glass) and more so in quartz, since all tetrahedra are partially or completely bound to each other by the common oxygen atoms. It will be shown later on that in practically all orthosilicates the internal vibrations of the individual tetrahedra do not appear. It is therefore more correct to consider the vibrations of the entire set of tetrahedra simultaneously. Since, however, many authors still use the concepts of isolated SiO_4 tetrahedra, referring, for example, to experimental data on the vibrational spectra of silicane ethers, we shall again return to a more detailed analysis of this problem.

The experimental results of Kujumzelis and Langenberg were essentially confirmed by Norris,[39] whose work we shall not discuss here.

There are many very interesting and extensively publicized investigations on the effect that annealing of glass has on its structure. There are only two papers by Prod'homme specially devoted to the study of glass of complex composition by the Raman scattering method. A description of the results of analogous experiments with fused quartz is found in one of the papers of Kujumzelis.[37]

According to Prod'homme, the hardening of the specimen leads to a unique effect in the Raman spectra, namely, to a certain broadening of all bands. No frequency shifts or redistribution of intensities among the bands is observed. Yet Kujumzelis[37] did not observe any changes in the spectrum of fused quartz at all. Prod'homme tends to attribute his results to a reduction in the degree of ordering of the SiO_4 tetrahedra in the hardened specimens.

Since the literature on the application of the Raman scattering method to the study of this problem is skimpy and since the results obtained

are to some extent contradictory, we confine ourselves to these remarks.

B. Two-Component Silicate Glass

The undisputed shortcoming of the papers discussed above is the not too felicitous choice of glass types, of too complex compositions. It is true that this circumstance did not prevent establishing the general character of scattering in glass-like substances, but it has undoubtedly made it difficult to determine their structure. A favorable exception if the investigation of two-component sodium and lead-silicate glass by M. F. Vuks and V. A. Ioffe.[44] Sodium glass was represented by seven specimens, Na: 23, 27, 33.3, 37.6, 42.9, 45.5, and 50. The lead-silicate group, in view of technological difficulties, comprised only three specimens, Pb: 35, 40, and 50.*

This research has established that the spectra of sodium-silicate glass contain four bands each, while those of lead-silicate glass have two bands. As the silicon content increases, the frequencies and intensities of the bands change systematically. In the limit, at the maximum silicon content in the glass, the frequencies of several of the bands approach the values characteristic of amorphous and crystalline quartz. One of the bands with a frequency of approximately 1090 cm^{-1}, is particularly intense in the spectra of the sodium-silicate specimens that are rich in sodium oxide. Apparently no analogue can be found for it in the quartz spectrum. The bands of lead-silicate glass are shifted towards the lower frequencies, relative to those of solium-silicate glass. M. F. Vuks and V. A. Ioffe correlate this to the replacement of the relatively light sodium atoms by the considerably heavier lead atoms.

To judge more reliably the possible structural changes in sodium-silicate glass, Vuks and Ioffe plotted the frequency variations of the four observed spectrum bands as functions of the composition of the glass (Fig. 1). Before discussing this plot, we make the following comment. In the investigation of various properties of multi-component systems it is assumed that these properties vary with the composition continuously and monotonically, so long as no new components appear or no old components vanish in the system. It is natural to think that deviations from the monotonic variation in the character of the Raman spectra (frequency, intensity, and polarization of the bands) can also serve, for some compositions, as a criterion of certain structural transformations in the system. Naturally, further data are necessary to establish accurately what transformations do, indeed, take place.

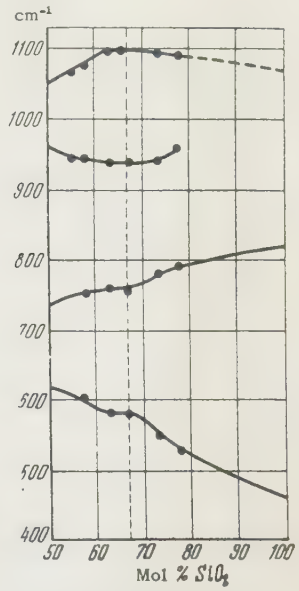


FIG. 1. Plot of the frequencies of certain bands in the spectra of sodium-silicate glass vs. composition.[44]

It is seen from Fig. 1 that as composition of the glass approaches that of the sodium-bisilicate type ($Na_2O \cdot 2SiO_2$), the frequency variation for all bands without exception, is no longer monotonic with the composition. This, apparently, is evidence of a structural transformation in the bisilicate region. Vuks and Ioffe believe that the bisilicate composition corresponds to the formation of a chemical compound in the glass. Incidentally, later one of these authors later refuted such an interpretation of the results.[45]

The idea of investigating sufficiently simple systems was further developed by E. F. Gross and V. A. Kolesova.[46, 47] They investigated sodium and potassium-silicate glass and also one sample of lithium glass and fused quartz. Their research has established that the frequency and intensity of the spectrum bands vary with the alkali content in the glass. For the sake of clarity, all their data are shown schematically in Fig. 2, while the frequency dependencies are shown in Fig. 3. Both figures are taken from references 46 and 47. It is evident from these data that the variations in the spectra are essentially monotonic. Only in compositions approaching the metasilicate do certain frequencies vary sharply. The specific frequency variation observed by Vuks and Ioffe[44] in the region of the sodium bisilicate is thus unconfirmed.

Comparison shows great similarity between the spectra of sodium-silicate and potassium-silicate glass, and also with the spectra of one

*Here and throughout the numbers denote the molecular percentages of the metal oxide.

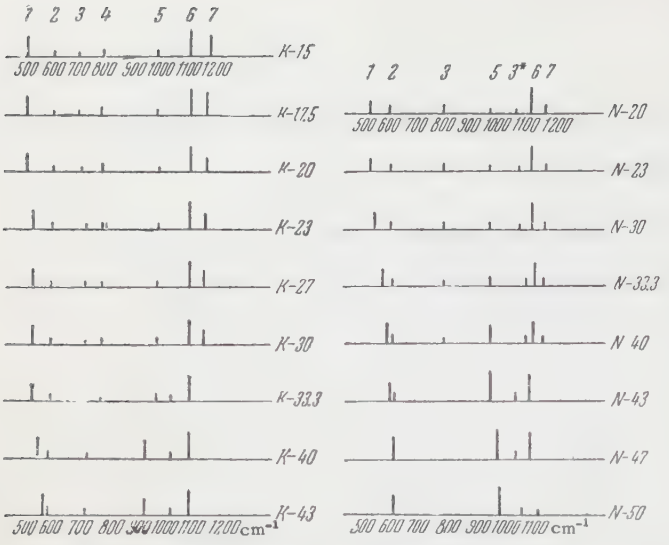


FIG. 2. Schematic representation of the spectra of sodium and potassium silicate glass.[46]

type of lithium glass. Yet lithium, sodium, and potassium differ quite substantially in their atomic weights. This means that the cations of the alkali metals form strong ionic bonds with the oxygen, and consequently their vibrations are inactive in Raman scattering. In other words, Raman spectra of alkali-silicate glass contain only the vibrations of the silicon-oxygen skeleton. This is exactly the interpretation proposed for this fact by Gross and Kolesova.[46, 47]

Incidentally, in the later one of these two papers,[47] Gross and Kolesova indicate that the spectra of the two specimens of Na-20 and K-20 glass display weak smeared bands near ~ 300 cm^{-1}. This can be ascribed to vibrations of the Na–O and K–O bonds, respectively. Since, however, a strong background of parasitic scattering appears when this region of the spectrum is investigated by the ordinary photographic procedure,[47] these results may be in doubt.

Even if the vibrations of the Me_{alk}–O bonds are actually inactive, this gives no grounds whatsoever for generalizing the results to include any other cation. Furthermore, investigation of glass containing different cations, which will be discussed later, proves without doubt the ability of many cations to become incorporated into the silicon-oxygen grid through partially covalent bonds.

Thus, the work of Gross and Kolesova leads to the important conclusion that the Raman scattering, at least in the case of alkaline-silicate glass, displays essentially the vibrations of the silicon-oxygen skeleton. This conclusion, reinforced by further research made by other authors,[21, 24, 25] has served as the base for a theoretical calculation of the spectral frequencies of certain silicates and glasses of known composition.

The most general point of view, which is consistently developed by Gross and Kolesova, is that to each band in the spectra of the investigated glass there is a corresponding band (analogue) in the spectrum of fused quartz. Since, as already noted, the spectra of glass display gradual variations with the composition, this might appear to prove the existence of a continuous structural transition from fused quartz to glass. Thus in the opinion of the Gross and Kolesova[46, 47] glass is, from the structural point of view, a spatially deformed silicate lattice with partially broken Si-O bonds (owing to the oxygen atoms in the oxide of the metal), in which cations are statistically distributed. In the limit, when the glass approaches the metasilicate in composition, such a lattice is gradually transformed into an infinite chain. This is just the point of view of Zachariasen,[14] which was later developed by his successors[48] and applied to complicated glass. This concept is particularly clearly formulated by Gross and Kolesova.[47]

In all the above investigations the spectra were recorded photographically. The authors could therefore study only the frequency variation, and instead of accurately measuring the intensities they confined themselves to semi-quantitative estimates. The situation was particularly unfavorable in connection with polarization measurements. Until recently the literature contained only two references to investigations in this direction. One[20] deals with research on polarization spectra of two random specimens of complex glass, and its results are therefore of value from the procedural point of view rather than from the scientific one. The second work[49] is devoted to an investigation of fused quartz, and will be referred to repeatedly.

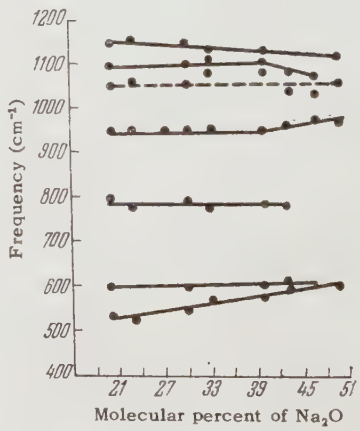


FIG. 3. Dependence of certain bands in the spectra of sodium-silicate glass on their composition.[46,47]

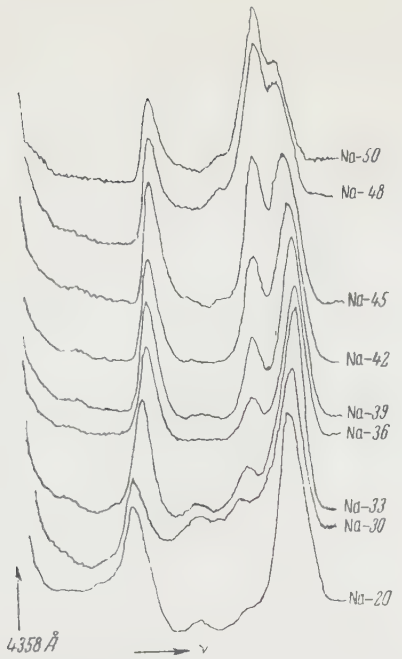


FIG. 4. Photoelectric records of spectra of several sodium-silicate types of glass.[25]

Only with the development of the photoelectric method of recording spectra did it become possible to fill the gap in the field of polarization measurements. At first these new experimental procedures* were applied to investigations of sodium-silicate and potassium-silicate glass.[21, 24, 25]

The first of these systems, which was studied in particular detail, was represented by 22 specimens, starting with glass (Na-12) and ending with metasilicate (Na-50). The composition was varied in steps of one to three molecular percent. In addition to these systems, spectra of lead metasilicate (Pb-50) and fused quartz were investigated.

The spectra of several sodium-silicate and of two potassium-silicate types of glass, (obtained in natural light), taken for references 21, 24, and 25, are shown in Figs. 4 and 5. The details characteristic of the scattering are clearly seen. In spectra of alkali-poor specimens, and also in complex lead glass or fused quartz,[36-38] a continuous Raman spectrum is seen adjacent to the excited line. This apparently does not appear at once, if the series of spectra is examined from the metasilicate side, but only starts with glass containing approximately 30% of alkali. Then, as the glass becomes poorer in alkali, the spectrum becomes more intense, but nevertheless of lower intensity than the analogous scattering in fused quartz. This is illustrated in Fig. 6, which shows

*All polarization measurements reported here and throughout were carried out with an illuminating beam of large aperture.

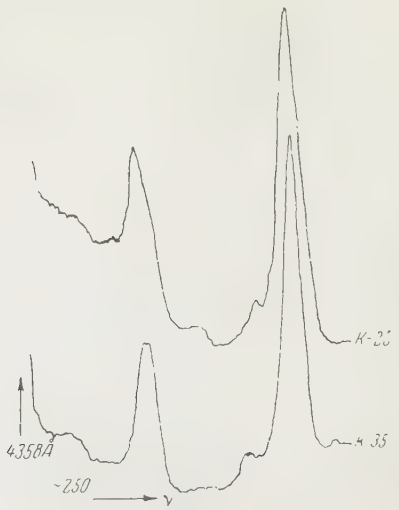


FIG. 5. Photoelectric records of spectra of two potassium-silicate glass specimens.[25]

the spectra of specimens of fused quartz and K-15 glass of equal size, obtained under identical conditions. Thus, an impression is gained that the continuous Raman spectrum is a specific feature of the spatial coupling of the SiO_4 tetrahedra.

It is also seen in Figs. 4 and 5 that a small diffused maximum is observed at low frequencies (approximately 250 to 300 cm^{-1}), particularly in spectra of potassium glass. Since the intensity of this maximum does not change noticeably with varying cation concentration,[25] it is apparently wrong to attribute it to vibrations of the Me—O bond, as has been done by Gross and Kolesova. It will be shown below that the polarization measurements confirm the correctness of these conclusions.

We now call attention to another important feature of the scattering spectra of specimens with low alkali contents, as compared with spectra of fused quartz, namely the very high intensity of the 1096 and 1170 cm^{-1} bands. It exceeds in intensity the "analogue band" of fused quartz by a factor of approximately 25. In our opinion[25] this shows that this group of bands in the spectra of fused quartz and glass is connected with different forms

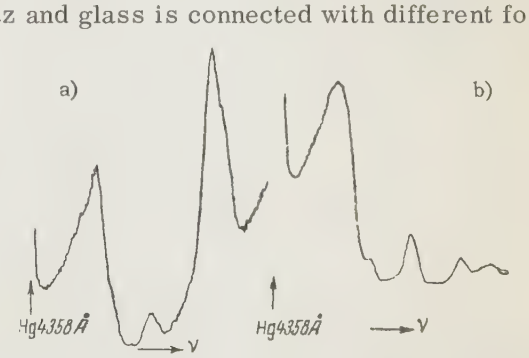


FIG. 6. Comparison of spectra a) of fused K-15 glass and b) quartz, obtained under identical conditions.

FIG. 7. Photoelectric recording of polarization spectrum of fused quartz: a) intense component, b) weak component.[25]

of vibrations. Consequently, it is apparently incorrect to consider these bands to be analogues. The same ratio of intensities of the considered bands is observed also for glass of the sodium-silicate system, for example for Na-12 glass.

The behavior of the band near 800 cm^{-1} is unique. Its intensity is sufficiently large in fused-quartz spectrum; in two-component glass the intensity diminishes and continues to diminish systematically the richer the glass is in alkali. This band disappears entirely from the spectra of glass that is close to metasilicate in composition. Analogous observations were made by Gross and Kolesova.[46]

The polarization spectrum of fused quartz, shown in Fig. 7, is characterized by a strongly polarized continuous background of Raman scattering, and also by two polarized bands at 495 and 610 cm^{-1} and by depolarized bands near 800, 1130,

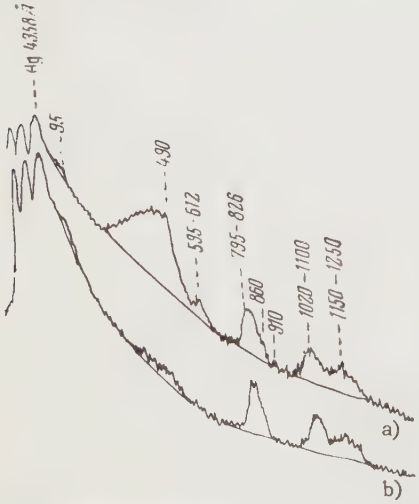


FIG. 8. Photometric curves of polarization spectrum of fused quartz: a) intense component, b) weak component.[49]

and 1230 cm^{-1}. These results are in good agreement with those previously obtained by Harrand.[49] For comparison, we include Harrand's photometric curves in Fig. 8. We note, however, that unlike Harrand's data, no weak polarized satellite of the band near 800 cm^{-1} has been observed in reference 25. This is evidently due to the great width of the spectral interval covered by the photoelectric apparatus of reference 25.

Polarized spectra of two-component silicate glass have much in common with the spectrum of fused quartz. Here, too, we observe a polarized continuous scattering background, a polarized band near 550 cm^{-1}, and a depolarized band near 800 cm^{-1}. In potassium glass a depolarized band is seen near to 250 cm^{-1} in the low-frequency region. The fact that this band is depolarized, along with the above-mentioned absence of noticeable variations in its intensity with the cation content, is convincing refutation of its belonging to vibrations of the K—O bond. It is more likely that this band is related, for example, to deformation vibrations of the silicate skeleton.

There is one important detail in the polarization spectra of all specimens, without exception, namely that the bands of highest frequency ($\sim$ 1096 to 1170 cm^{-1}), which are the "analogues" of the bands near 1130 and 1230 cm^{-1} in fused quartz, are strongly polarized. Comparision of the latest results on the relative intensities of the considered bands with those previously described leaves no doubt that this group of bands, in two-component glass and in fused quartz, is connected with different forms of vibrations. This in turn makes quite doubtful the correctness of the ideas concerning the gradual transition from the fused-quartz structure to that of two-component glass, as developed by Gross and Kolesova.[46, 47] It is apparently more correct to assume that the transition from quartz even to glass with the minimum cation content is accompanied by a radical and thorough structural rearrangement. Were it otherwise, the polarization of the bands could not change in an abrupt manner. Starting with the ideas of gradual structural changes, it is also difficult to imagine that the polarization of the above bands increases in the transition from quartz to glass, because the lattice of glass with defects in the form of disorderly distributed metal cations (the point of view of Gross and Kolesova[46, 47]) is even less regular and symmetrical than the lattice of fused quartz. It will be shown below that a theoretical calculation confirms the conclusions that follow from this experimental investigation.

Also of interest is the behavior of the band near

970 cm^{-1}. In samples of glass containing up to 28 or 30% alkali, this band, being quite weak and diffused, is apparently depolarized. In the bisilicate region ($Na_2O \cdot 2SiO_2$), the band becomes much stronger and pronouncedly polarized. It is hardly possible to explain this essentially-qualitative jump other than to assume a superposition of bands belonging to two different structures, of which one (corresponding to the polarized band) appears in compositions close to the bisilicate. Henceforth, as the alkali content in the glass increases, the new polarized band becomes reinforced even more sharply, becoming the most intense band in the spectrum of potassium metasilicate.

One of the characteristic features of the spectrum of sodium metasilicate is that it is richer in bands as compared with the spectra of glass of the preceding compositions. Polarized light discloses another feature. It turns out that in this case, unlike in the preceding types of glass, the bands with frequencies near 970 and 1070 cm^{-1} are polarized differently: the latter band must be considered depolarized or at least only lightly polarized (Fig. 9).

The spectrum of lead metasilicate differs in many respects from what would appear in the spectrum of the related sodium metasilicate. First, it is simpler than the spectrum of sodium metasilicate. Furthermore, instead of the clearly resolved and intense bands near 970 and 1070 cm^{-1}, which characterize the latter spectrum, we observe here a very broad and intense band with traces of a fine structure and with a maximum near 960 cm^{-1}. It is all completely polarized. In addition, this glass is characterized by an intense continuous Raman scattering, which is entirely absent from the spectrum of sodium metasilicate. All this gives grounds for assuming that the two types of glass the metasilicates of lead and sodium, have substantially different structures.

The measurement data given in reference 25 have been used by the authors to plot the variation of the frequencies of the most intense bands in the spectra of the sodium silicate glass with composition (Fig. 10). Comparing the plot of Fig. 10 with that previously obtained by Vuks and Ioffe (Fig. 1) it is easy to establish their practically complete identity. This also refutes the data (Fig. 3) of Gross and Kolesova[46, 47] given above.

Thus, measurements of frequency and of the degree of depolarization in spectra of sodium-silicate glass are also evidence of a structural rearrangement in the bisilicate region. An analogous conclusion was reached previously by V. A. Florinskaya[50] as the result of an investigation of infrared spectra of these types of glass. It is interesting to

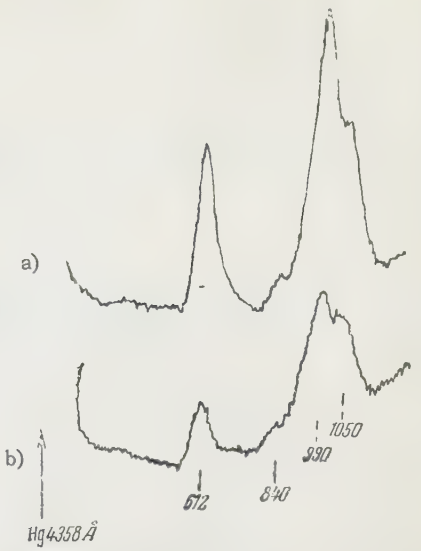


FIG. 9. Photoelectric recording of polarization spectrum of Na-50 (metasilicate) glass: a) intense component, b) weak component.[25]

note that this is also confirmed by certain other measurements, for example by measurements of viscosity of glass.[51] This is why the suggestion by M. F. Vuks that glass contains a chemical compound of sodium bisilicate, corresponding to the layered structure of glass in this region, is probably correct and is undoubtedly worthy of attention.

Conclusions. The most general result of the investigations of Raman spectra of complex silicate glass is a more or less reliable establishment of the fact that the glass contains various large groupings of atoms and anions and, above all, interlinked SiO_4 tetrahedra. There are two points of view concerning the character of the linkage of these tetrahedra in simple alkali-silicate glass. According to the first one, the interlinking of the tetrahedra obeys statistical laws, by virtue of which glass of any composition (with the exception of metasilicate) does not constitute a chemical compound. The authors of the second point of view presuppose the

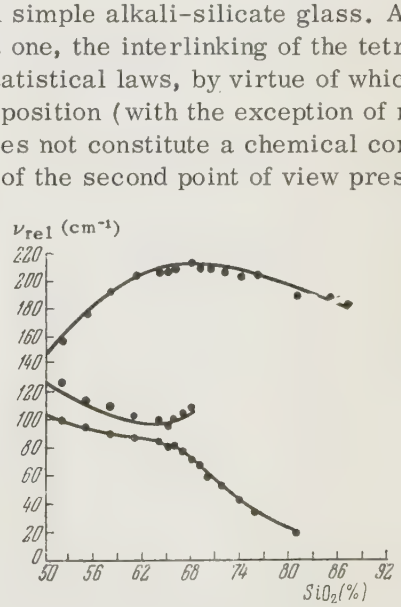


FIG. 10. Dependence of the frequencies of certain bands in sodium-silicate spectra on composition.[25]

structural inhomogeneity of glass, whereby glass of any composition is considered as a superposition of spatial-lattice, layered, and chain-type structural elements. In particular, they assign the chemical formula $Na_2O \cdot 2SiO_2$ to glass of the sodium-bisilicate composition, thus explaining its layered structure.

C. Investigation of the Influence of Different Chemical Elements on the Structure of Silicate Glass [26,27]

The influence of chemical elements on the structure of silicate glass has been the subject of various investigations, but unfortunately by indirect methods as a rule. Much of the work performed and described in the literature is therefore based on purely conjectural, frequently speculative ideas concerning the possible influence of various properties of its component parts on the process of vitrification and, in the final analysis, on the atomic-molecular "architecture" of glass.

There is hardly another physical or physical-chemical method that can give more valuable information on this topic than the method based on the study of the vibrational spectra. Everything discussed in this section has been obtained by means of Raman scattering of light.

We chose[26, 27] as objects for our investigations mixed metasilicates of the $NaO \cdot MeO \cdot 2SiO_2$ type and orthosilicates of the $Na_2O \cdot Me_2O_3 \cdot 2SiO_2$ and $2Na_2O \cdot MeO_2 \cdot 2SiO_2$ types, for the case of trivalent and tetravalent Me respectively. The latter, as can be seen, include the oxides of sodium and of various other elements in that total molecular amount at which not a single SiO_4 tetrahedron is bound to another common oxygen atom.

The choice of this composition of glass is far from arbitrary. It is easy to show in this case that to any imaginable structure of glass there can be ascribed definite spectral distinguishing features. Let us consider this problem in detail, using orthosilicates as an example.

Assume that the element Me is present in this glass in the form of cations. Then the spectrum must show the SiO_4 tetrahedra, since the introduction of sodium oxide into the glass, as was shown in references 46 and 25, breaks the Si—O—Si bonds. Only one polarized band in the spectrum corresponds to the case of regular tetrahedra. If the SiO_4 tetrahedra are deformed in the Coulomb field of the cations so strongly that they have lost all their symmetry elements, this is indicated in general by the absence of polarized bands.

In that case when the Me—O bond has an essentially covalent character, we can assume that the oxide of the element Me enters into the general lattice with Na_2O and SiO_2. We expect that this must display a continuous background of polar-

TABLE III. Data on the frequencies and polarizations of bands in spectra of mixed meta and orthosilicate glass[27]

Serial number	Glass	Frequencies (in cm^{-1}) and polarization of the bands
1	$40Na_2O \cdot 10BeO \cdot 50SiO_2$	Cs (sp), 620 (sp), 1065 (pp)
2	$*Na_2O \cdot MgO \cdot 2SiO_2$	Cs (sp), 635 (sp), 1010 (pp)
3	$*Na_2O \cdot CaO \cdot SiO_2$	335 (d), 500 (d), 620 (sp), 985 (pp) 1060 (pp)
4	$*Na_2O \cdot SrO \cdot 2SiO_2$	335 (d), 620 (sp), 980 (pp), 1050 (pp)
5	$*Na_2O \cdot BaO \cdot 2SiO_2$	310 (d), 440–460 (d), 600 (sp), 962 (pp), 1075 (pp)
6	$*Na_2O \cdot ZnO \cdot 2SiO_2$	Cs (sp), 625 (sp), 1045 (pp)
7	$*Na_2O \cdot CdO \cdot 2SiO_2$	Cs with boundary at 475–540 (sp), 625 (sp), 985 (pp)
8	$*Na_2O \cdot PbO \cdot 2SiO_2$	Cs with boundary at ~530 (sp), 650 (pp), 962 (pp)
9	$40Na_2O \cdot 10Al_2O_3 \cdot 50SiO_2$	Cs with boundary at ~565 (sp), 1065 (pp)
10	$**Na_2O \cdot Bi_2O \cdot 2SiO_2$	Cs with boundary at ~445 (sp), 980 (pp)
11	$**Na_2O \cdot B_2O_3 \cdot 2SiO_2$	Cs with boundary at ~510 (sp), 630 (sp), 730 (?), 1067 (pp)
12	$**Na_2O \cdot TiO_2 \cdot 2SiO_2$	Cs with boundary at ~300 (sp), broad band at 345 (d), 700 (sp), 875 (sp), 1015 (d)
13	$**2Na_2O \cdot GeO_2 \cdot 2SiO_2$	295 (d), 580 (sp), 855 (pp), 965 (pp), 1092 (pp)
14	$40Na_2O \cdot 10ZrO_2 \cdot 50SiO_2$	Cs (sp), 605 (sp), 930 (pp), 1075 (pp)
15	$**2PbO \cdot SiO_2$	Cs (sp), with boundary at ~500, 925 (pp)

Symbols: sp – strongly polarized, pp – partially polarized, d – depolarized, cs – continous scattering.

Note: Glass marked with one or two asterisks has (by analysis) a composition quite close to meta- and ortho-silicate, respectively. All other types of glass have not been analyzed and their composition is assumed to be in accordance with their synthesis.

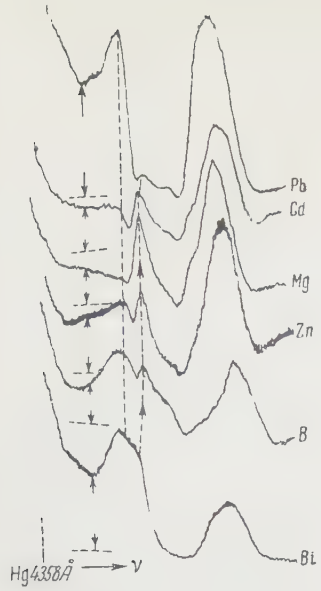


FIG. 11. Photoelectric recordings of spectra of several mixed meta-and orthosilicate types of glass. The dashed line shows the level of the background near the exciting line, and the arrows indicate the continuous scattering.[26,27]

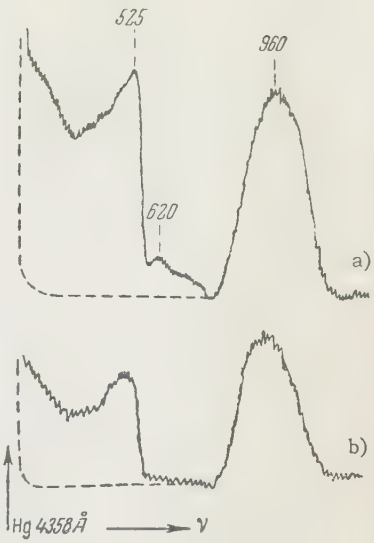


FIG. 12. Photoelectric recording of polarization spectrum of mixed metasilicate glass, containing PbO as the third component: a) intense component, b) weak component.[26,27]

ized Raman scattering adjacent to the exciting line, analogous to the scattering in fused quartz and in two-component alkali-poor glass, i.e., with strongly-developed spatial linking of the tetrahedra.[25] It will be shown below that this is fully confirmed experimentally. In addition, the bands of the spectra will probably be particularly broad in this case.

It is not excluded, for example, that the oxides of the element Me, particularly if they are glass-forming themselves, will not enter at all into the common lattice, and will more likely tend to form their own lattice. A superposition of the spectra of the glass-forming oxide and of the metasilicate $Na_2O{:}SiO_2$ must then be observed. The latter has been thoroughly investigated by us experimentally[24,25] and by B.I. Stepanov and A. M. Prima[52-55] theoretically, and has a sufficiently characteristic form. Naturally, intermediate structures are also possible. The spectra of metasilicate glass can be interpreted in approximately the same manner.

In references 26 and 27 we reported on a study of the influence of all the most important readily-available elements (Be, Mg, Ca, Ba, Si, Zn, Cd, Pb, Al, Bi, B, Ti, Ge, Zr). Three of these—Be, Al and Zr—were investigated in glass somewhat different from that described above (with a lower relative content of oxides of these elements). It was possible to produce and investigate lead glass in the form of a simple orthosilicate $2PbO{\cdot}SiO_2$.

Even the results of the investigation of lead orthosilicate show that the character of its spectrum is actually compatible with the assumption of the existence of a spatial lattice made up of SiO_4 tetrahedra, linked to each other by the lead atoms. It is remarkable that the spectrum of lead metasilicate, as already noted, also exhibits features of such links, manifesting itself in particular in the presence of a continuous background of polarized Raman scattering. It must therefore be assumed that in this glass all the tetrahedra are linked, at least partially, through the lead atoms.

Further investigation of the above-mentioned mixed meta and ortho-silicates has made it possible to subdivide all elements into three groups, in accordance with their effects on the character of the spectra.

Table III lists the measured frequencies in spectra of various types of glass. Spectra of certain representatives of the first group, the most extensive one, are shown in Fig. 11. The polarization spectrum of one type of glass of this group is shown in Fig. 12. These spectra are characterized by a more or less intense continuous background of polarized scattering, with a small peak near its boundary. This peak retains its position (near 625 cm^{-1}) approximately in all glass types of this group. The spectra also contain very broad high-frequency bands with a maximum near 980 cm^{-1}. The widths of the bands, and particularly the presence of a continuous scattering background, shows convincingly in our opinion, a partial spatial linkage of most SiO_4 tetrahedra via the Me atoms. Aluminum, bismuth, lead, and boron have a particularly strong tendency to penetrate the structural glass lattice. Judging from the presence of a band near 625 cm^{-1} which is characteristic of the metasilicate,[46,25] it appears that a small portion of the Na_2O and SiO_2 molecules, which do not enter into a common lattice with the element Me, forms structures that are close to metasilicate chains. The number of chain structures differs with different types of glass. A more or less stronger tendency to enter

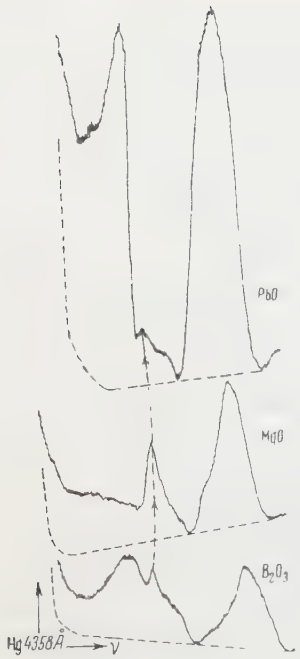


FIG. 13. Comparison of the spectra of glass containing PbO, MgO, and B_2O_3. The spectra have been recorded under identical conditions.[27]

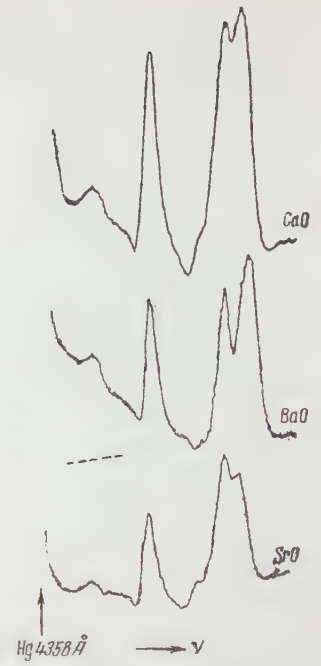


FIG. 14. Photoelectric recording of spectra of mixed metasilicate glass, containing CaO, BaO, and SrO.[26,27]

into the structural glass lattice is displayed also by beryllium, magnesium, zinc, cadmium and zirconium.

When discussing the problem of the tendency of chemical elements to enter into the general structural lattice, we must not lose sight of the following. The intensity of the Raman spectra, as noted earlier, is directly related to the character of the bonds. The more covalent the bonds, the more intense the spectra, in general. Consequently, operating with intensities in the scattering spectrum, we are essentially speaking, characterizing the "covalent contribution" of the chemical element in the structural lattice. This is why an estimate of the tendency of the chemical elements to enter into the structural lattice should be made taking into account the general intensity of the spectrum.

In connection with this, it is of undoubted interest to compare spectra obtained under perfectly identical conditions, of different types of glass. In view of the fact that many specimens were different in dimensions and quality, we could make this comparison[26, 27] only for three types of glass, containing PbO, MgO, and B_2O_3. The spectra are shown for comparison in Fig. 13. We see that the most intense spectrum is that of lead glass, and the less intense is that of boron glass. The magnesium glass occupies an intermediate place.

One might think that the high intensity of the spectrum of lead glass, lightly tinted yellow, is due to the relative nearness of the exciting line (λ = 4358Å) to the corresponding absorption band (resonance phenomenon). This doubt is however completely dispelled in the case of uncolored magnesium glass. Apparently, it is the degree of covalence of the Me—O bonds which manifests itself. This means that the Mg—O bond in the glass is more covalent than the B—O bond.

At first glance this conclusion is paradoxical. Calculation of the degree of ionization of the bonds, using the electro-negative potential scale of Pauling, gives the opposite results (see Table I). One must note, however, that this scale is, strictly speaking, applicable only to isolated bonds. In such a complex condensed system as glass, the situation may change.

In our opinion, a convincing illustration of this fact is germanium glass GeO_2. The results of the investigation of certain types of germanium-containing glass will be reported below. We note here only that its spectrum is several times more intense than the spectrum of glass-forming silicon, although both types of glass have structures which are identical in all respects. In addition, according to calculations, the Ge—O bonds and the Se—O bonds have an equal degree of covalence which amounts to approximately 50%.

Thus, estimates of the degree of ionization (covalence) of the bonds in glass, based on the use of the ordinary scale of electro-negative potential of the elements, have a roughly approximate character and do not take into account the mutual influence between the different structural formations in the glass. We note that similar considerations have already been stated in the literature.[11]

The second group, according to spectroscopic

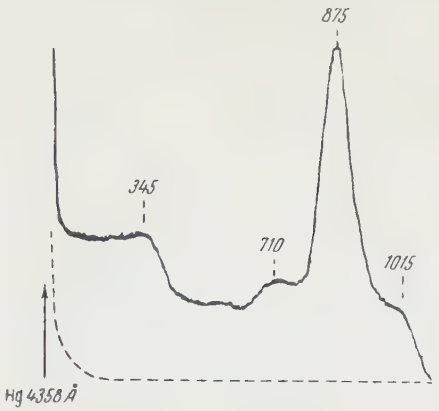


FIG. 15. Photoelectric recording of the spectrum of the glass $2Na_2O \cdot TiO_2 \cdot 2SiO_2$. The dashed line shows the level of the background.[27]

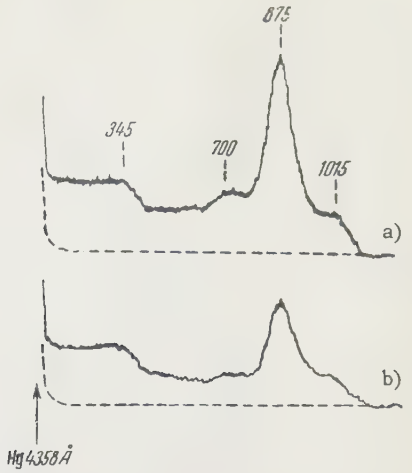


FIG. 16. Photoelectric recording of the polarization spectrum of the glass $2Na_2O \cdot TiO_2 \cdot 2SiO_2$: a) intense component, b) weak component.[27]

investigations, includes the glass types that contain CaO, BaO, and SrO. Their spectra are shown in Fig. 14. They are characterized by somewhat narrower bands, by the absence of a continuous background of polarized scattering (instead of which we see two comparatively weak depolarized maxima, ~ 320 and ~ 470 cm^{-1}), by a well resolved structure of the high frequency band, and by somewhat different intensity distributions. In their appearance and in all details, with the exception of the maxima at the ~320 and 470 cm^{-1}, all the spectra obtained recall those of the sodium-silicate glass, which is close in composition to the metasilicate. The similarity is particularly pronounced in the case of glass containing SrO.

Since these types of glass are metasilicates with respect to their total content of Na_2O and MeO, the facts noted above lead to the conclusion that the metals Ca, Ba, and Sr enter into the glass in the form of cations (like Na and K), and consequently the spectra of these types of glass manifest vibrations of the silicon-oxygen metasilicate chains.

The third group contains only one single type of glass with composition $2Na_2O.TiO_2.2SiO_2$. The spectrum of this glass, shown in Fig. 15, is most unique. Unlike the preceding spectra, the bands near 625 and 980 cm^{-1} have disappeared from this spectrum. They are replaced by a very intense and polarized band near 875 cm^{-1} (see Fig. 16). Furthermore, the spectrum also displays a continuous background of polarized scattering with a strongly shifted boundary (apparently to approximately 750 to 800 cm^{-1}). Pronounced against the background of this scattering are a depolarized band (with boundary near 345 cm^{-1}) and a polarized band (near 700 cm^{-1}). Adjacent on the high-frequency side to the band near 875 cm^{-1} is a weak depolarized satellite with a frequency of 1015 cm^{-1}.

The impression is gained that in this case weak vibrations of the free SiO_4 tetrahedra and of the complex spatial lattice $SiO_2.TiO_2$ appear in the spectrum. In this connection, the band near 875 cm^{-1} is interpreted by us[26,27] as belonging to the completely symmetrical vibration of the SiO_4 tetrahedra. The 1015 cm^{-1} band is ascribed to one of the triply-degenerate anti-symmetrical vibrations, and the very broad band with a high-frequency boundary at nearly 345 cm^{-1} is attributed to another triply-degenerate and to a doubly-degenerate vibration. This interpretation of the bands of the tetrahedron vibrations is confirmed by theoretical calculations.[52-55] It is also in good agreement with the results of an investigation on a large number of other tetrahedral molecules[56] and with data on polarization measurements.[26,27]

Another fact worthy of attention is that the boundary of the continuous scattering background in titanium glass, unlike that in other silicate glass, shifts sharply towards the higher frequencies. To explain this shift (with allowance for the opposing effect of the mass of the titanium atom), it is necessary to assume that the Ti—O bond is considerably more rigid than the Si—O bond. It is not excluded that the strengthening of the Ti—O bond is due to the three d-electrons, which are the external electrons of the titanium atom and which participate in the formation of the bond.

For the sake of clarity, Figs. 17 to 19 show schematically the structures of several types of glass as derived from the analysis of the spectroscopic data.

It is of interest to compare the conclusions concerning the tendency of different elements to enter into the structural lattice drawn from spectroscopic investigations and from other investigations. It is commonly accepted that aluminum has

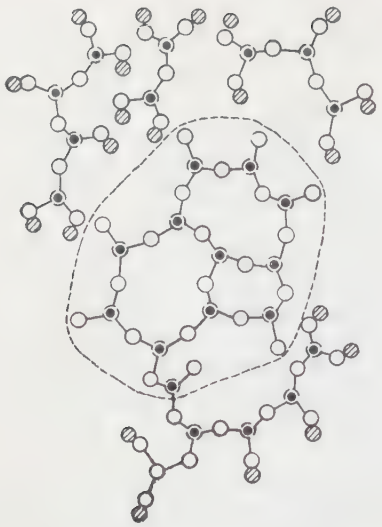


FIG. 17. Schematic representation of the structure of two-component alkali-silicate glass. Symbols: o) O, ⊙) Si, ◉) Na. The dashed line separates the spatial lattice region. Outside the dashed region are traces of layers made of SiO_4 tetrahedra.[26,27]

a very pronounced such tendency. Data on Raman spectra, as already noted above, confirm this fully. The results of the investigation of the thermal expansion of various types of glass, made by Stanworth,[57] are compatible with the assumption that magnesium and zinc can enter into the structural lattice, while calcium and barium cannot. These facts are also in good agreement with the spectroscopic data. According to Fajans and Kreidl,[58] and also according to a later paper by Stanworth,[60] the tendency to enter into the structural lattice is a characteristic of lead and bismuth atoms. Recently it has even been possible to melt "silicate" glass with the total molecular content of Bi_2O_3 and PbO reaching 95%.[60] Such types of glass essentially

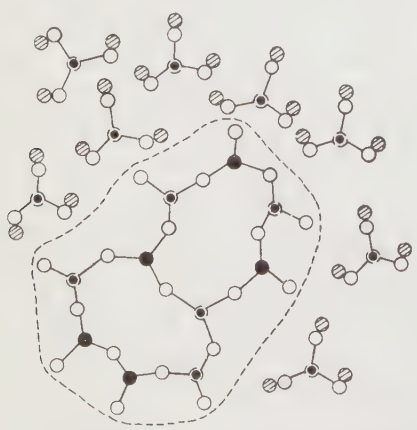


FIG. 19. Schematic representation of the structure of a three-component glass containing TiO_2 as a third component. Symbols: o) O, ●) Si, ⊙) Ti, ◉) Na. The dashed line separates the spatial-lattice region composed of SiO_4 and TiO_4 tetrahedra. Outside the dashed region are the free SiO_4 tetrahedra.[26,27]

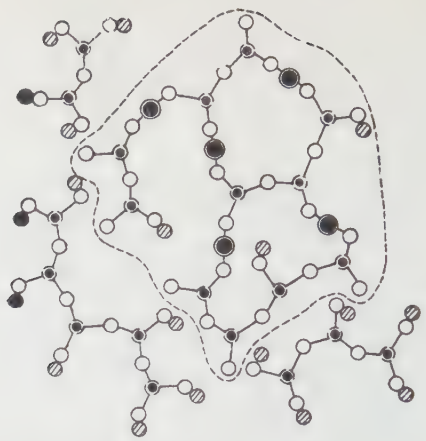


FIG. 18. Schematic representation of the structure of three-component glass containing PbO as a third component. Symbols: o) O, ●) Si, ⊙) Pb. The dashed line separates the spatial-lattice region in which the SiO_2 tetrahedra are partially bound through the lead atoms. Chains of SiO_4 are shown outside the dashed region.[26,27]

cannot even be called silicate glass. These facts are also in good agreement with the results of the investigation of the Raman spectra.

Thus, the interpretation we have proposed[26,27] for the Raman spectra of mixed meta- and orthosilicates, can be considered as fully acceptable and can serve as a basis for further experimental investigations of similar type.

It is also of interest to compare the results of spectroscopic investigation with the conclusions concerning the structure of various types of glass, formulated by Dietzel.[7,8] We recall that according to Dietzel the "strong cations" (such as B^{+3}, Bi^{+3}, Ti^{+4}, Ge^{+4} etc.) should not show a strong tendency to enter the silicon-oxygen lattice, but should tend, on the contrary, to produce their own surroundings of oxygen anions.* In this sense they compete with the cations of silicon. Spectroscopic data, at least for those types of glass investigated in references 26 and 27, do not confirm this conclusion fully. The picture obtained is much more complicated than that which follows from Dietzel's conjectures.

Let us examine this problem in somewhat greater detail.

According to Dietzel, there should exist in three-component sodium-borosilicate glass isolated boron-oxygen and silicon-oxygen lattices with uniformly distributed sodium cations. The glass we investigated had (according to analysis) the composition of an orthosilicate, $Na_2O.B_2O_3.2SiO_2$. Consequently, the oxygen atoms of the silicon-oxygen lattice should be surrounded by twice as many Na^+

*Again we recall that according to Dietzel the tendency towards formation of a general structural lattice is completely determined by the electrostatic interaction forces.

cations as the oxygen atoms of the boron-oxygen lattice. If this is so, the spectrum should contain a superposition of the vibrations of Na-25 borate and silicate glasses.

Fig. 20 shows, along with the spectrum of the three-component glass, the spectra of the two-component glasses Na-20 (silicate) and Na-20 (borate). The small variation in the compositions of these types of glass is insignificant. As can be seen, no superposition of the spectra is observed. Consequently, the Dietzel rule is not obeyed in this case.

Matters are even worse with titanium glass. Here, as already noted, it is reasonable to assume the existence of a complex lattice of the $SiO_2.TiO_2$ type and of free tetrahedra, in which the negative charges of the oxygen atoms are compensated by the charges of the sodium cations. Yet, according to Dietzel, the spectrum should have displayed at least the bands of glass having the Na-40 composition.

Neither does Dietzel's theory allow for the behavior of such metals as bismuth, lead, cadmium, zinc, and several others. This is particularly clearly seen in the case of glass containing Bi_2O_3. No features of the Na-25 composition are observed in the spectrum of this glass (see Fig. 11).

An example where the Dietzel rules are exactly obeyed is the mixed orthosilicate of germanium, $2Na_2O.GeO_2.2SiO_2$, the spectrum of which is shown in Fig. 21. From the requirement of uniform distribution of sodium cations (the Dietzel concept) it follows that the spectrum of this orthosilicate should display the bands of the Na-40 two-component silicate and germanium glasses. This is indeed observed experimentally, as shown directly for the first of the two types of glass by Fig. 21. Thus, an analysis of the spectroscopic data makes it possible to refine and to correct the widespread ideas concerning the structure of glass.

In connection with the foregoing material, we also comment on certain incorrect, but widely held interpretations of the infrared spectra of various crystalline silicates.

It is known that silicates can exist in the form of various structures. Some of these form a spatial lattice. A typical representative of such silicates is quartz. Others form layers and chains. There are, finally, silicates to which one ascribes the so-called "island" structure. These are crystalline analogues of orthosilicate glass. Schafer, Matossi, and Wirz[41] have investigated, among others, the infrared spectra of silicates with "island" structure: Be_2SiO_4—phenacite, (CaOTi) SiO_4—titanite, (Zn, Mg) SiO_4—troostite,

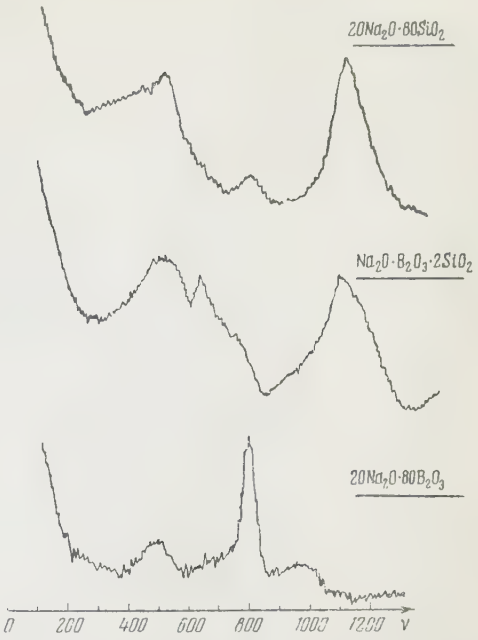


FIG. 20. Comparison of spectra of mixed orthosilicate $Na_2O{\cdot}B_2O_3{\cdot}2SiO_2$, Na-20, silicate glass, and Na-20 borate glass.

and others. This reference also cites literature data on zircon $ZrSiO_4$ and willemite Zn_2SiO_4. The analysis of the spectra has made it possible for the authors of reference 41 to attribute four bands belonging to these silicates to vibrations of the free tetrahedra of SiO_4. These data are then used to interpret the vibrational spectra of glass.

As follows from the results of even a roughly approximate study of Raman scattering, no free tetrahedra of SiO_4 appear in orthosilicate glass. Consequently, the identification of the corresponding bands in the spectra of crystalline silicates, as proposed by Schafer, Matossi and Wirz, is not convincing.

We confine ourselves for the present to the

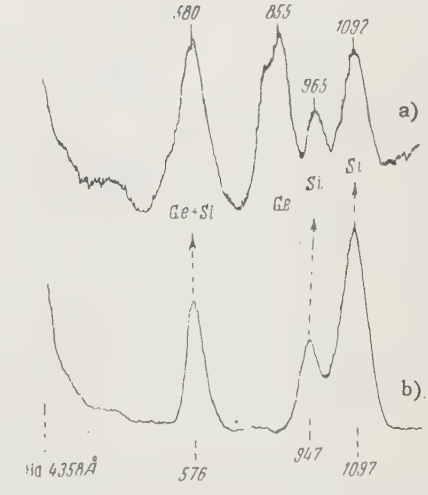


FIG. 21. Comparison of spectra of the mixed orthosilicate $2Na_2O{\cdot}GeO_2{\cdot}2SiO_2$ (a) and Na-40 silicate glass (b).

description of the character of the spectra of various types of glass and to an interpretation of their structure. It is difficult to explain at this stage the different behavior of the chemical elements without additional experimental material. In particular, it is exceedingly important and interesting to carry out analogous investigations on other systems, for example phosphate and borate. To show how difficult and involved is the problem of the influence of chemical elements on the structure of glass, we call attention again to certain of the types of glass investigated in references 26 and 27. As already noted, the oxides of titanium and germanium influence differently the structure of three-component glass. Yet titanium and germanium have identical valence and do not differ greatly in the values of the ionic radii. Boron and bismuth, on the other hand, have substantially different ionic radii, but apparently behave in glass in an analogous manner.

Conclusions. The investigation of Raman spectra of mixed meta and orthosilicates indicates the diverse roles of the chemical elements in the vitrification process. The spectra are separated into three groups by their character. Each spectrum is amenable to a definite interpretation and leads to conclusions on the structure of the investigated glass. There is full agreement between the behavior of the chemical elements as established by spectroscopic means and by other indirect methods.

D. Spectra of Liquid Silicates (Orthosilicic-Acid Ethers Considered as Analogues of Silicate Glass

More than 20 years ago Singer and Weiler[61] and Singer[62] undertook to interpret spectra of certain silicates and silicate glass on the basis of an experimental study of liquid silicates (orthosilicane ethers), i.e., compounds having a general formula $Si(OC_mH_n)_4$. Recently an analogous attempt was repeated by A. N. Lazarev.[63] The importance of such comparisons is as follows. As already noted, the principal structural element in glass is the SiO_4 tetrahedron. According to x-ray structural investigations of monomeric methyl ether $Si(OCH_3)_4$, carried out by Eulitz,[64] a SiO_4 group appears in the substance, with the same symmetry and the same dimensions as in the crystalline quartz. Thus, by studying the vibrational spectra of the monomeric ether and its condensation products it would appear possible to separate the frequencies of the internal vibrations of the SiO_4 tetrahedron and follow their variation as a function of the degree of polymerization of the $Si(OCH_3)_4$ molecules and then apply the data obtained to the interpretation of the spectra of glass.

This, indeed, is the procedure followed in references 61 and 62. Analyzing their own experimental data, they found it possible to ascribe to the vibrations of the SiO_4 tetrahedral group in the spectrum of monomeric methyl ether the following frequencies (in cm^{-1}): 1205, 1192, 1170 (ν_1), 1110-1080 (ν_2), 842 (ν_3), and 642 (ν_4). Since the vibrations ν_1 and ν_2 are three-fold degenerate, their splitting into several components, in the opinion of Singer and Weiler, just implies the removal of the degeneracy. According to Weiler[62] the polarization measurements do not contradict the proposed interpretation.

As the degree of polymerization of the ethers increases (upon transition from the monomeric to the dimeric, trimeric and decameric ether) the bands ν_1 and ν_2 practically retain their position in the spectrum while the bands ν_3 and ν_4 shift towards lower vibration frequencies (this is particularly true of the latter, which reaches rapidly a limiting value of approximately 520 cm^{-1}). The limiting values of the frequencies in the spectrum of the investigated ethers are compared by Singer and Weiler with the frequencies of the bands in the spectrum of crystalline quartz.

A. N. Lazarev's investigation[63] of methyl and ethyl ethers of silicate acids are based on similar considerations. The difference consists only in that the treatment of the results obtained is somewhat different: Lazarev compares the limiting values of the frequencies of the silicon skeleton of the ethers with the frequencies of alkali-silicate glass, having a composition close to that of the metasilicate, and made up of infinitely long silicon-oxygen chains, rather than using Weiler's comparison with the frequencies of crystallne quartz. Lazarev establishes here an almost complete correspondence between the spectra of both objects.

We doubt seriously that the spectra of monomeric ethers of silicic acid can display discrete vibrations of the SiO_4 tetrahedra. It is much more correct to propose the possibility of interaction between the vibrations of the Si—O and C—O bonds. In the interpretation of the bands it then becomes necessary to take into account a different, lower than Td, symmetry of the ether molecule as a whole. A solution of the mechanical problem of the vibrations of the $Si(OCH_3)_4$ molecules, carried out by Iguchi,[65] has shown that in this case four polarized bands should appear in a spectrum among the skeleton vibrations.

In order to verify this premise, Lazarev, Tulub, et al.[66] undertook a thorough investigation of the

TABLE IV. Data on the frequencies, intensities, and polarizations of bands in spectra of different silicane ethers[66]

$Si(OCH_3)_4$	$OSi_2(OCH_3)_6$	$Si(OC_2H_5)_4$	$OSi_2(OC_2H_5)_6$	$O_2Si_3(OC_2H_5)_8$
412 (1; 0.80)	445 (1.5;)			
	577 (10; 0.05)		614 (2; 0.13)	604 (1.5; 0.18)
640 (10; 0.03)		656 (6; 0.05)		640 (0)
	692 (0)		685 (0)	703 (0.5; ?)
844 (2; 0.80)	808	792		
	838 (5.5; 0.50)	822 (3; 0.62)	800 (1.5; 0.31)	796 (2; 0.34)
1082	1090	939 (3; 0.53)	935 (1; 0.78)	940 (1; 0.80)
1111 (5; 0.20)	1126 (6; 0.40)	1096 (8; 0.27)	1095 (2.5; 0.27)	1096 (2.5; 0.22)
1186 (4; 0.32)	1190 (4.5; 0.30)	1173 (4.5; 0.15)	1170 (0)	1170 (0)
1269 (1.5; pol.)	1267 (2; pol.)	1205 (2; 0.74)		
		1302 (4.5; 0.77)	1296 (1.5; 0.70)	1295 (2; 0.48)
1371 (1; pol.)	1375 (1; pol.)			
1466 (5.5; 0.80)	1466 (8; 0.80)	1460 (10; 0.83)	1456 (4; 0.70)	1456 (4; 0.55)

Symbols: The first number in the parenthesis indicates the relative intensity of the band, the second its degree of depolarization. Weak lines, observed photographically, are assigned everywhere a zero intensity value.

polarization spectra of monomeric methyl and ethyl ethers, and also of certain condensation products of these ethers. The degrees of depolarization were measured by the photoelectric method. The data obtained are listed in Table IV, and some of the recordings are shown in Fig. 22.

We see that the spectrum of methyl ether, in contradiction with Weiler's experimental data,[62] contains not one but several polarized lines. This is in good agreement with the theoretical calculations of Iguchi.[65] An approximately analogous picture is observed in the case of other ethers. This proves that the four bands in the spectrum of the $Si(OCH_3)_4$ molecules cannot be attributed to vibrations of SiO_4 tetrahedra and raises in general serious doubts concerning the possibility and advisability of the very attempts of paralleling the spectra of silicane ethers and those of silicate glass for the purpose of determining the structure of glass.

3. THEORETICAL INTERPRETATION OF RAMAN SPECTRA OF CERTAIN CRYSTALS. COMPARISON WITH EXPERIMENTAL DATA FOR SILICATE AND GERMANIUM GLASS

Attempts at a theoretical interpretation of vibrational spectra of crystalline silicates have been undertaken many times. Thus, for example, Saksena[67] calculated the frequencies in the spectrum of α-quartz, and Barriol[68] did the same for β-quartz. These calculations, however, are rough and the results are presented in a very inconvenient form, which makes comparison with experiment difficult. Matossi[69] derives expressions for the vibration frequencies of the pyroxene chain of SiO_4 tetrahedra, of tetrahedra linked in pairs, and of quartz. It is difficult to judge the accuracy of these expressions. We can only say that, in particular for the pyroxene chain, they are known to be incorrect, since Matossi employs incorrect values for the cosine of the Si–O–Si angles.

The application of the results of calculations of the spectra of crystalline silicates to the case of glass is fraught with certain specific difficulties. It is necessary to know how the tetrahedra are linked with each other in the glass and what role is played by various chemical elements in the structure of the glass. Direct experiment cannot answer the first question. It is more correct to seek a confirmation in the results of the theoretical calculations of the frequencies, based on a sufficiently reasonable structural model. The different views

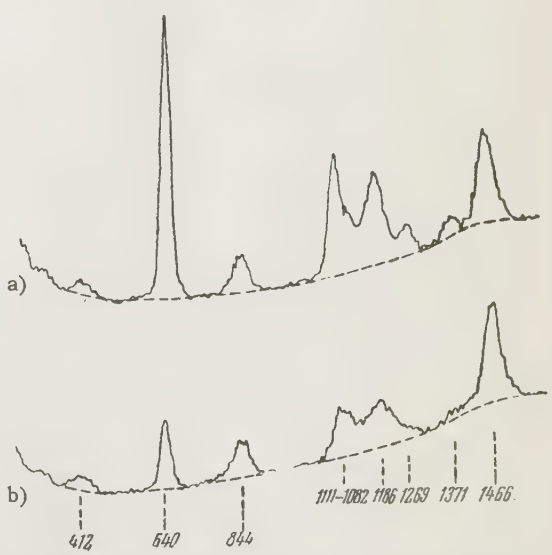


FIG. 22. Photoelectric recording of the polarization spectrum of $Si(OCH_3)_4$: a) intense component; b) weak component.[66]

on the character of the linkages between the SiO_4 tetrahedra in alkali-silicate glass have already been discussed.

The situation is somewhat simpler when it comes to the influence of chemical elements on the structure of the glass. The available experimental data make it possible to judge quite reliably what elements are involved and to what extent they enter into the silicon-oxygen lattice of the glass.[26,27] This in turn determines the applicability of the theoretical calculations and points the way to further refinement of these calculations.

Until recently there were practically no investigations in which the results of the theoretical calculations of the frequencies in the spectra of crystalline silicates were used to determine the structure of glass. The only attempt of this kind was made by Bobovich, Girin et al.[21] These authors presuppose the existence of lattice structures, layer structures, and chain structures in alkali-silicate glass. Applying the expressions derived by Matossi[69] for the vibrations of a chain to certain bands in the spectrum, they succeed in interpreting many important experimental facts. Nevertheless the results obtained are far from conclusive, since the expressions for the frequency are themselves doubtful.

Only the insufficient development of the theory

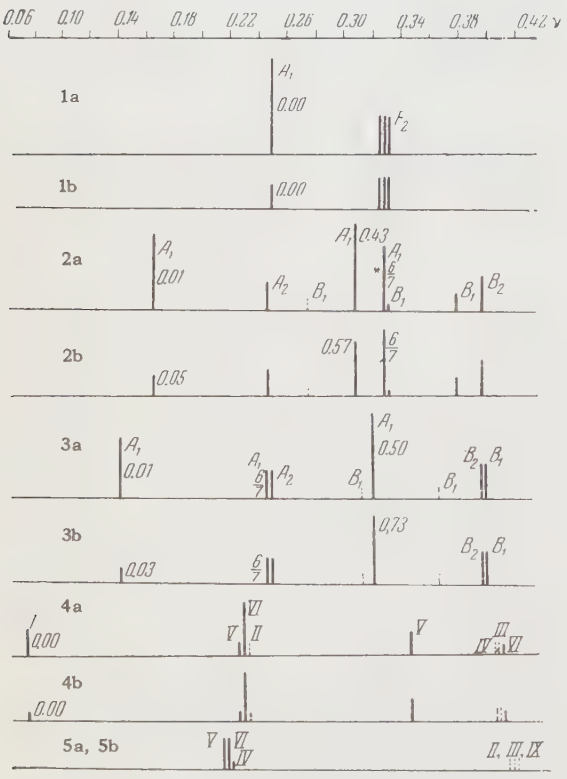


FIG. 23. Theoretical Raman spectra: 1) SiO_4 tetrahedron, 2) chain, 3) layer, 4) β-quartz, 5) β-cristobalite; a) with $\partial\alpha_{\parallel}/\partial q = 0$, b) with $\partial\alpha_{\perp}/\partial q = 0$ (references 52-55).

of the problem can explain the fact that for a long time the authors of many papers[70-74] have guided themselves in the interpretation of experimental data by the definitely incorrect hypothesis, first stated by Schafer, Matossi and Wirtz,[41] that the spectra of glass retain the vibration forms of isolated SiO_4 tetrahedra. To some extent this problem has been treated by us earlier. A correct critique of the hypotheses of Schafer, Matossi and Wirtz and of their successors is contained in the work by V. A. Kolesova.[75]

A very important step forward in the interpretation of the vibrational spectra of crystalline silicates and of glass was taken recently by B. I. Stepanov and A. M. Prim.[52-55] They have considered the vibrations of β-quartz and of cristobalite, free SiO_4 tetrahedra, the pyroxene chain, and an infinite layer of SiO_4 tetrahedra. The last case was analyzed first. All the calculations were made in two versions, a rigorous one and one with reasonable simplifications, using a method previously developed by M. A. El'yashevich and B. I. Stepanov as applied to vibrations of molecules. In addition to the frequencies, they also calculated the relative intensities of the vibrational bands, and also the states of polarization of Raman bands. For this purpose they used an additivity scheme, proposed by M. V. Vol'kenshtein, and took into account the variations in the polarizability components in direction of and transverse to the bonds.

Without going into details, let us list the principal results of the simplified version of the calculations, when the effect of the deformation vibrations is excluded. In all cases, with the exception of quartz and cristobalites, two elastic constants are introduced. These results are then compared with the experimental material.

For convenience, some of these results, pertaining only to Raman spectra, are gathered in the form shown schematically in Fig. 23. The following conventional notation is used. The dotted lines indicate the bands that are forbidden in the used approximation, while double lines signify the natural bands. The letters A, B, and F with indices and Roman numbers indicate the type of symmetry of the vibrations, while the numbers represent the states of polarization of the bands. The vibrations denoted by A_2, B_1, B_2 and F, as well as by V and VI, are depolarized. The relative intensities of the Raman bands depend strongly on the ratio of the quantities $\partial\alpha_{\perp}/\partial q$ and $\partial\alpha_{\parallel}/\partial q$. Figure 23 shows the results of calculations for two extreme cases: $\partial\alpha_{\perp}/\partial q = 0$, $\partial\alpha_{\parallel}/\partial q \neq 0$ and $\partial\alpha_{\perp}/\partial q \neq 0$, $\partial\alpha_{\parallel}/\partial q = 0$. When $\alpha_{\perp}/\partial q = \partial\alpha_{\parallel}/\partial q$, the selection rules allow only the fully-symmetric vibrations. As

follows from the available experimental data for crystals, it is apparently the intermediate case that is realized in practice. The asterisks indicate the random degenerate vibrations, for which the degeneracy is removed if the different Si—O bonds are characterized by two elastic constants of different values.

It is seen from Fig. 23 that each investigated structure is characterized by a specific spectrum. In particular, in quartz and cristobalite, unlike in the chain and in the layer, the depolarized bands lie in the region of the highest vibration frequencies. The spectrum of β-cristobalites appears simpler than the spectrum of β-quartz; this, of course, is caused by its greater symmetry. One might think that this feature alone could identify the particular structure. Unfortunately, the lack of sufficient experimental material does not for the time being permit a consistent verification of the theoretical results for crystals. We note, however, that the data contained in reference 76 appear not to confirm the relative simplicity of the spectrum of β-cristobalite.

To compare the above schematic spectra of crystalline silicates with real spectra of glass it is necessary, as already noted, to make a definite assumption concerning the structure of glass. Stepanov and Prima[52-55] started out with the assumption that fused quartz, sodium bisilicate, and sodium metasilicate represent, as regards structure, arbitrarily oriented minute crystals of quartz and two-dimensional and undimensional silicate crystals, respectively. This has made it possible to calculate the intensities and the degree of depolarization of the bands, using the same rules as apply in the case of ordinary systems with freely oriented particles.

Naturally, the model chosen for glass is too primitive from the physical point of view. If we start out with the crystalline concepts, it is more correct, in any case, to have the term crystallite signify a near-order region that has no phase boundaries, a fact to which we have already called attention. The choice on the part of Stepanov and Prima[52-55] is justified, however, by considerations of simplicity of calculation. A refinement of the glass model, in the sense formulated above, could hardly add anything new to the results of the calculations.

Attention must be called to still another circumstance. Strictly speaking, the experimental data, at least for fused quartz, should be properly compared with the theoretical ones for β-quartz and not for α-quartz, which is a quartz modification that is stable in the region of ordinary temperatures. Stepanov and Prima[52-55] did not carry out the corresponding calculations. One can expect, however, [77, 78] that allowance for this circumstance will not introduce substantial corrections. In any case, the corrections will not exceed the errors associated with the approximate character of the calculations. Having made these remarks, let us turn to a comparison of the theoretical and experimental data on the spectra of glass.

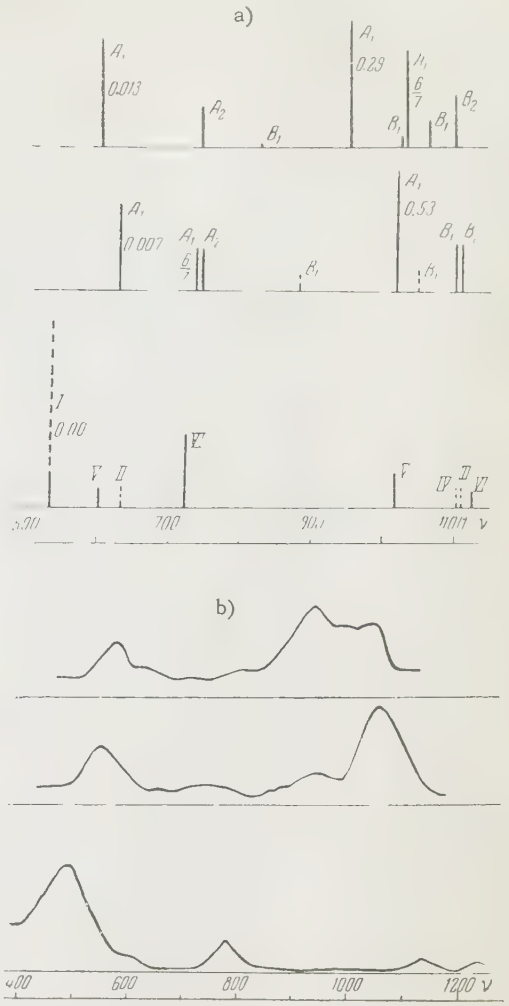


FIG. 24. Theoretical (a) and experimental (b) Raman spectra of various types of glass (upper row — metasilicate glass, middle row — bisilicate glass, lower row — fused quartz).[52-55]

As we did previously, we now present the results of the calculations schematically, showing alongside, also in schematic form, the experimentally-obtained spectra (Fig. 24). In the calculations we have assumed for the elastic constants values that give their best agreement with the experimental data for sodium metasilicate.[25] In addition, for components of the polarizability in direction of and transverse to the bonds, we assumed the relation $\partial\alpha_{\perp}/\partial q = 1/4\ \partial\alpha_{\parallel}/\partial q$, which is frequently satisfied in vibrations of different molecules.

As can be seen from Fig. 24, the calculated and observed spectra are in agreement with respect to all the substantial features. We note above all that results of calculations and their comparison with experimental data do not contradict the hypothesis of the layered structure of glass with the composition of sodium bisilicate ($Na_2O.2SiO_2$). To obtain additional proof of the correctness of this hypothesis, we have applied elsewhere[79] the results of the theoretical calculations of Stepanov and Prim to the spectrum of sodium bigermanate ($Na_2O.2GeO_2$). In addition, we took into account very close analogy in the structure of the crystalline germanium dioxide and that of α-quartz, thanks to which the silicate and germanium glass can be considered as formed by isotopic oxides of the same type. The frequencies of the spectrum of sodium bigermanate, calculated on this basis, are in good agreement with the observed ones. This proves the existence of the layered structure of sodium bisilicate, inasmuch as the equations for the vibrations used in the calculations were indeed derived under the assumption of a layered linkage of the SiO_4 tetrahedra.

Another fact deserving of attention is that, according to calculation, the spectrum of fused quartz, unlike the bisilicate spectrum, to which a layered structure is ascribed, contains a low-intensity depolarized vibration, V, in the region of approximately 1050 to 1200 cm^{-1}. Consequently, our statement[21, 25] that the above bands in the spectra of fused quartz and of glass are not analogues, is fully corroborated by calculation.

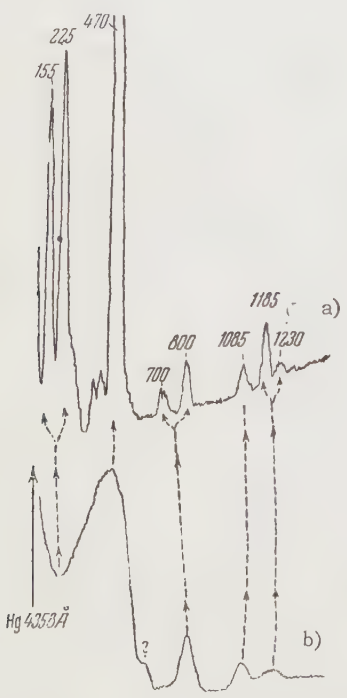


FIG. 25. Photoelectric recordings of spectra: a) crystalline quartz, b) fused quartz. The latter specimen was cut perpendicular to the crystallographic axis.

A theoretical interpretation is similarly found for several other experimental facts, observed by the authors:[24, 25] particularly, that the band with frequency near 1070 cm^{-1} in the spectrum of sodium metasilicate is depolarized or at least only partially polarized; that the bands with frequencies from approximately 525 to 600 cm^{-1} in the spectra of all types of glass are very intense and polarized, etc. Only in some details of little importance does a theoretical analysis refine and correct our conclusions in references 21 and 25.

In connection with the theoretical analysis of the spectra of silicate glass, we must dwell specially on the problem of the structure of the simplest type of glass—glass-like silica (fused quartz). The following problem is frequently raised in experimental investigations: what does the structure of glass-like silica approximate more closely, crystalline quartz or cristobalite? Since the Raman spectrum of cristobalite is still unknown, only one possibility remains, that of comparing the spectra of glass-like silica and crystalline quartz. As early as in 1928,[34] E. F. Gross and M. F. Romanova established the great similarity between the spectra of the two materials. Considerably later, in 1952, E. F. Gross and V. A. Kolesova confirmed these results and noted that the "broad bands in the spectrum of glass-like quartz are located in approximately the same positions as the sharp pronounced lines in the spectrum of crystalline quartz."[80] Other authors, citing analogous data, do not specially emphasize the similarity, leaving the readers to judge for themselves.

Thus it appears that one can conclude that there is a similarity in the structural lattices of glass-like silicon and of crystalline quartz. In other words, one can assume that the SiO_4 tetrahedra in glass-like silicon make with each other on the average an angle of 147°, which is characteristic of quartz.

However, one year later (1953) E. F. Gross and V. A. Kolesova stated that there is no similarity in the spectra and consequently in the structure, of crystalline and fused quartz. They stated that the latter more likely has the structure of cristobalite, but do not cite thereby any new spectroscopic data.[47] As arguments in favor of this assumption, Gross and Kolesova cite the results of an investigation on glass-like silica by roentgenoscopic methods and by the method of neutron diffraction, the results of

TABLE V. Comparison of data on frequencies in the spectra of crystalline and fused quartz obtained by various authors

Crystalline quartz (frequencies in cm^{-1})			Fused quartz (frequencies in cm^{-1})				
Gross and Romanova[34]	Kujumzelis[37]	Bobovich and Tulub	Gross and Romanova[34]	Kujumzelis[37]	Krishnan[82]	Harrand[49]	Bobovich and Tulub*
125	127	155 (s)			30—120	95	
207	209	225 (s)	215				
264	266		265				
	295				285		
320			325	230—450		275—500	
358	357	350 (w)	365		370		
405	398	415 (w)	445		430		
	410						
463	465	470 (vs)					
503			500	500	495	500	500 (vs)
526	530						
585	570—600						
633			625	607	635	595—612	595—600 (w)
694		700 (w)	665		660		
746	740		741				
800	803	800 (w)	790—830	780—840	775—805	795—826	800 (m)
944	860—950				810—845	860	
1021	1025				825—940	910	
	1063						
1075	1083	1085 (w)	1020—1070	1030—1090	1022—1098	1020—1100	1065 (w)
	1100						
1163	1162	1185 (w)					
1220	1235	1230 (w)	1180—1230	1160—1230	1140—1245	1150—1250	1200 (w)
	1400—1600						

*The frequencies measured photoelectrically at the maximum of the bands. Literature data on the frequencies are found in reference 25.

Symbols: vs – very strong, s– strong, m – moderate, w – weak.

surface examinations, and also the fact that cristobalite crystals are invariably precipitated during its devitrification.

In connection with the latter fact, we note the following. The very fact that one type of crystal or another is precipitated is not convincing proof of a corresponding structure of the original, non-crystallized body. To judge the structure of glass reliably, it is necessary to follow its "thermal life" step by step, as is done, for example, in the papers of V. A. Florinskaya and her associates.[50] At this point, E. F. Gross is again in contradiction with his own discussion of the glassy state, where he stated: "In fact, it is impossible to consider crystallized glass, in which crystals of various silicate compounds are precipitated, as a certain analogue of the structure of the true homogeneous glass. It would not occur to anyone to consider a liquid in which crystals are precipitated as a system similar in structure to the true liquid."[81]

It appears to us that even raising the question of comparison of the vibrational spectra of fused quartz, crystalline quartz, and cristobalites, is rather meaningless. In fact, at ordinary temperatures fused quartz can contain only the α-modification of crystalline quartz or cristobalite. Their similarity is so great that one can hardly expect great spectroscopic differences. On the contrary, the β-modifications are substantially different. However, they are stable only at high temperatures.

Nevertheless we thought it advisable to compare the Raman spectra of glass-like silica and of crystalline quartz, obtained under identical conditions. These spectra are shown in Fig. 25. To evaluate their reliability, we selected data on both spectra from the best investigations.

A careful study of the data given in Table V, taken from various sources, permits the conclusion of a sufficiently good agreement as regards the principal lines and bands. On the other hand, we see also a close agreement between the spectra of glass-like silica and crystalline quartz. The existence of several weak lines, the very reliability of which can be doubted, in the spectrum of the crystalline quartz can hardly serve as an argument against this conclusion. Consequently, the experimental data on Raman spectra of glass-like silica and crystalline quartz are compatible with the assumption that their structural lattices are analogous.

The results of the theoretical calculations con-

firm well the conclusions of the experimental research. According to calculations, the spectrum of glass-like silicon is characterized by two sufficiently intense bands, one very intense and polarized at the low-frequency boundary of the investigated region, and another weaker one, depolarized in the central region. The remaining bands should be weak and depolarized. All this is, indeed, observed experimentally. The small discrepancy between the calculated and observed values of the frequencies in the range from 1180 to 1100 and from 1150 to 1250 cm^{-1} is explained by the fact that reference 25, on which the calculation is based, has incorrect values of the maximum frequencies of these bands. The use of exact experimental values (see Table V) leads to good agreement in the results. Were the glass-like silica closer in structure to cristobalite than according to the calculations, its spectrum would contain only one intense band.

It must be noted that in one substantial detail the experimental and theoretical data are still in disagreement. As was shown by Harrand[49] and by the authors,[25] a polarized band is observed in the spectrum of glass-like silica in the region around 600 cm^{-1}. The possibility of an experimental error must be excluded. Yet, according to calculations, this region should contain a depolarized band. One can therefore assume that the glass-like silica contains additional structural formations, although, judging from the intensities of the bands, the quartz-like formations predominate.

In connection with the discussion of the problem of the structure of glass-like silica, it is appropriate to cite certain results obtained by the method of infrared spectra. The advantage of this method, incidentally, consists of the relative ease with which minute crystals and powdered substances can be investigated. It is in exactly this form that cristobalite and tridymite can be prepared. That is why it has been possible recently to obtain and compare the spectra of all modifications of quartz.

Infrared spectra are usually studied in absorption or reflection. However, only the absorption spectra have a simple physical meaning. Generally speaking, the absorption spectra can be obtained from the reflection spectra through computation. However, the computational method, owing to certain arbitrariness in the initial assumptions, may lead to systematic errors in the results. This is why data on absorption spectra and conclusions on the structure of glass-like silica, made by several authors using the computational scheme[76] are always under doubt.

It is much more advisable to compare only transmission spectra or only reflection spectra. This is indeed the procedure used by V. A. Florinskaya and her associates.[72,150,83] According to their data, vitreous silicon is closer in its structure to crystalline quartz than to tridymite or cristobalite. This supports the conclusions based on an investigation of Raman spectra of certain of these bodies, and also the results of the theoretical calculations.

Conclusions. One can state that the theoretical calculations confirm the point of view that alkali-silicate glass is structurally inhomogeneous, as expressed in references 21, 24, and 25 on the basis of experimental investigation of their Raman spectra. The theoretical results are compatible with the assumption of a layered structure of glass with a composition of bisilicate. This is equivalent to saying that a chemical formula can be given for these types of glass. In accordance with the experimental data, calculations show that the glass-like silica is closer in structure to crystalline quartz than to cristobalite. From the very character of the calculation it follows that it is possible to consider various structural formations in glass as crystallites in the ordinary sense of the word.

[1] G. Tamman, Стеклообразное состояние (The Vitreous State), ONTI, 1935.

[2] P. P. Kobeko, Аморфные вещества (Amorphous Substances), Acad. Sci. Press, 1952.

[3] M. V. Vol'kenshtein and O. B. Ptitsyn, Dokl. Akad. Nauk SSSR **103**, 795 (1955).

[4] V. M. Goldschmidt, Trans. Far. Soc. **25**, 253 (1929).

[5] K. H. Sun and A. Silverman, J. Amer. Cer. Soc. **25**, 97 (1942).

[6] A. G. Pincus, Cer. Age **39**, 38 (1942).

[7] A. Dietzel, Z. Elektrochem. **29**, 537 (1941).

[8] A. Dietzel, Naturwiss. **48**, 9 (1942).

[9] L. Pauling, The Nature of the Chemical Bond, Oxford, 1940 (Russ. Transl. Goskhimizdat, 1947).

[10] G. Hägg, J. Chem. Phys. **3**, 42 (1935).

[11] H. Cole, J. Soc. Glass. Technol. **31**, No. 142, 114 (1947).

[12] B. Neumann, and H. Richter, Z. Elektrochem. **31**, 484 (1925).

[13] A. Winter-Klein, Verres et refractaires, No. 3, 147 (1955).

[14] W. H. Zachariasen, J. Amer. Chem. Soc. **54**, 3841 (1932).

[15] A. A. Lebedev, Тр. ГОИ (Trans. State Opt. Soc.) **2**, No. 10 (1931); **3** (1924).

[16] Строение стекла (The Structure of Glass, Transactions of the Conference on the Structure of Glass) Acad. Sci. Press, 1955 (Transl. by Consultants Bureau, NY).

[17] Vol'kenshtein, El'yashevich, and Stepanov, Колебания молекул (Vibrations of Molecules) vols. I and II. Gostekhizdat, 1949.

[18] G. Herzberg, Vibrational and Rotational Spectra of Monatomic Molecules, [Van Nostrand, prob. vol. 1, 1950] (Russ. Trans.) IL, 1949).

[19] K. Kohlrausch, Raman Spectra (Russ. Transl.) IL 1952, p. 32.

[20] D. H. Rank and A. E. Douglas, J. Opt. Soc. Am. **38**, 966 (1948).

[21] Bobovich, Girin, and Tulub, Dokl. Akad. Nauk SSSR **105**, 61 (1955).

[22] Ya. S. Bobovich and V. M. Pivovarov, Usp. Fiz. Nauk **60**, 689 (1956).

[23] Ya. S. Bobovich and D. B. Gurevich, J. Exptl. Theoret. Phys. (U.S.S.R.) **27**, 318 (1954).

[24] Ya. S. Bobovich and T. P. Tulub, J. Phys. Chem. (U.S.S.R.) **30**, 1679 (1956).

[25] Ya. S. Bobovich and T. P. Tulub, Оптика и спектроскопия (Optics and Spectroscopy) **2**, 174 (1957).

[26] Ya. S. Bobovich and T. P. Tulub, Paper at the XI Conference on Spectroscopy, Acad. Sci. Press (in press).

[27] Ya. S. Bobovich and T. P. Tulub, Оптика и спектроскопия (Optics and Spectroscopy) in press.

[28] B. A. Kiselev, Оптика и спектроскопия (Optics and Spectroscopy) **1**, 597 (1956).

[29] Ya. S. Bobovich and D. B. Gurevich, Dokl. Akad. Nauk SSSR **85**, 524 (1952).

[30] A. Hollaender and J. W. Williams, Phys. Rev. **34**, 380 (1929); **38**, 1739 (1931).

[31] C. V. Raman, Ind. J. Phys. **2**, 387 (1928).

[32] C. V. Raman and K. S. Krishnan, Ind. J. Phys. **2**, 398 (1928).

[33] P. Pringsheim and N. Rosen, Z. Physik **50**, 741 (1928).

[34] E. Gross and M. Romanova, Z. Physik **55**, 744 (1929).

[35] S. Bhagavantam, Ind. J. Phys. **6**, 1 (1931).

[36] Th. G. Kujumzelis, Z. Physik **97**, 561 (1935).

[37] Th. G. Kujumzelis, Z. Physik **100**, 221 (1936).

[38] R. Langenberg, Ann. Physik **28**, 104 (1937).

[39] R. Norris, Proc. Ind. Acad. Sci. **14A**, 178 (1941).

[40] W. L. Bragg, The Structure of Silicates (Russ. Transl.) ONTI 1934.

[41] C. Schäfer, F. Matossi, and K. Wirtz, Z. Physik **89**, 210 (1934).

[42] L. Prod'homme, Compt. rend. **233**, 303 (1951).

[43] L. Prod'homme, Verres et réfractaires, No. 6, 305 (1954).

[44] M. F. Vuks and V. A. Ioffe, Izv. Akad Nauk SSSR, OTN, No. 361 (1938); «Физ. хим. свойства тройной системы: окись натрия, окись свинца, кремнезем» (сборник) (Collection: The Physico-Chemical Properties of the Triple System Sodium Oxide, Lead Oxide, Silica) Acad. Sci. Press, 1949, p. 164.

[45] Строение стекла (The Structure of Glass; see reference 16), p. 327.

[46] E. F. Gross and V. A. Kolesova, J. Phys. Chem. (U.S.S.R.) **26**, 1673 (1952).

[47] Строение стекла (The Structure of Glass; see reference 16), p. 56.

[48] B. E. Warren and A. D. Loring, J. Amer. Cer. Soc. **18**, 269 (1935).

[49] M. Harrand, Compt. rend. **238**, 784 (1954).

[50] Строение стекла (The Structure of Glass; see reference 16), p. 70.

[51] E. Eipeltauer and G. Jangg, Kolloid Z. **142**, 77 (1955).

[52] B. I. Stepanov and A. M. Prima, Оптика и спектроскопия (Optics and Spectroscopy), in press.

[53] B. I. Stepanov and A. M. Prima, Оптика и спектроскопия (Optics and Spectroscopy), in press.

[54] A. M. Prima, Труды Института физ. и мат. (Trans. Inst. Phys. and Math.) No. 2, p. 124, Press of Acad. Sci. Belorussian S.S.R., Minsk, 1957).

[55] A. M. Prima, Paper delivered to XI Conference on Spectroscopy, Izv. Akad. Nauk SSSR (in press).

[56] K. Kohlrausch, Raman Spectra, (Russ. Transl.) IL, 1949, p. 134.

[57] J. E. Stanworth, J. Soc. Glass Technol. **30**, 54 (1946).

[58] K. Fajans and N. Kreidl, J. Amer. Cer. Soc. **31**, 105 (1948).

[59] J. E. Stanworth, J. Soc. Glass Technol. **32**, 154 (1948).

[60] S. M. Brekhovskikh, Стекло и керамика (Glass and Ceramics) No. 8, 1 (1957).

[61] R. Signer and J. Weiler, Helv. Chim. Acta **16**, 115 (1932).

[62] J. Weiler, Z. Physik **80**, 617 (1933).

[63] A. N. Lazarev, Физич. сборник (Physics Collection), vol I. No. III, p. 440. Press of the L'vov Univ. 1957. Paper delivered at X Conference on Spectroscopy.

[64] W. Eulitz, Z. Kristallogr. **80**, 204 (1931).

[65] K. Jguchi, J. phys. et radium **16**, 401 (1955).

[66] Lazarev, Tulub, and Bobovich, Оптика и спектроскопия (Optics and Spectroscopy) **4**, 417 (1958).

[67] B. D. Saksena, Proc. Ind. Acad. Sci. **12A**, 93 (1940).

[68] J. Barriol, J. phys. et radium **7** (8), 209 (1946).

[69] F. Matossi, J. Chem. Phys. **17**, 679 (1949).

[70] Ya. I. Gerlovin, Dokl. Akad. Nauk SSSR **38**, 136 (1943).

[71] N. Brügel, Z. Physik **128**, 255 (1950).

[72] V. A. Florinskaya and R. S. Pechenkina, Dokl.

Akad. Nauk SSSR **85**, 1265 (1952).
[73] V. A. Florinskaya and R. S. Pechenkina, Dokl. Akad. Nauk SSSR **89**, 37 (1953).
[74] V. A. Florinskaya, Dokl. Akad. Nauk SSSR **89**, 261 (1953).
[75] V. A. Kolesova, J. Exptl. Theoret. Phys. (U.S.S.R.) **26**, 124 (1954).
[76] I. Simon and H. O. McMahon, J. Chem. Phys. **21**, 23 (1953).
[77] T. M. K. Nedungadi, Proc. Ind. Acad. Sci. **11A**, 86 (1940).
[78] P. K. Narayanaswamy, Proc. Ind. Acad. Sci. **28A**, 415 (1948).
[79] Ya. S. Bobovich and T. P. Tulub, **Оптика и спектроскопия** (Optics and Spectroscopy), in press.
[80] E. F. Gross and V. A. Kolesova, Article in Memorial Collection to S. I. Vavilov, Acad. Sci. Press, 1952, p. 231.
[81] **Строение стекла** (The Structure of Glass; see reference 16), pp. 334-335.
[82] H. S. Krishnan, Proc. Ind. Acad. Sci. **37A**, 377 (1958).
[83] V. A. Florinskaya and I. A. Sevchenko, Dokl. Akad. Nauk. SSSR **109**, 1115 (1956).

Translated by L. E. Bergstein and J. G. Adashko

TEMPERATURE DEPENDENCE OF THE INTERNAL FRICTION OF METALS AND ALLOYS

V. S. POSTNIKOV

Usp. Fiz. Nauk **66**, 43-77 (September, 1958)

I. INTRODUCTION

As is well known, oscillations induced in a solid are damped out rapidly even in the absence of external resistance. It is customary to explain such damping by the presence of "internal friction forces" or simply "internal friction" within the solid itself. This term must be taken to mean the ability of the solid to convert, in an irreversible way, the energy of mechanical oscillations into heat. These "forces of internal friction" amount up to 70% (or more) of all the resistive forces that act on a vibrating solid in an actual situation.

The magnitude of internal friction may be defined in various ways. In the simplest and most direct way it may be expressed by means of a quantity δ defined, if the internal friction is not great, as the ratio

$$\delta = \frac{\Delta U}{U}, \qquad (1)$$

where ΔU is the energy dissipated per cycle, while U is the energy of the oscillations.

At room temperature for materials with low internal friction $\delta \sim 10^{-7}$, while for materials with high internal friction the values of δ may reach 0.1 or more. The value of δ is determined not only by the nature of the materials, but depends also, as we shall see later, on many other factors and above all on the amplitude of the oscillations. In the case of well tempered metals and of alloys, an amplitude dependence of the magnitude of internal friction determined, for example, by the method of low frequency rotational oscillations, can be observed for angles of twist exceeding 10^{-5} radian. The appreciable dependence of the magnitude of δ on the amplitude of oscillation should be taken into account in comparing the results of measurements obtained by different investigators. A direct measurement of δ, possible only in the case when $\delta \geq 0.1$, has been carried out by Hopkinson[1] and Föppl.[2]

Since in the deformation of a real solid one always observes a phase lag of the deformation behind the stress, i.e., mechanical hysteresis is present, it is also possible to adopt as a measure of internal friction the phase shift angle φ.[3,4,5,6] A direct determination of φ is possible only when $\tan\varphi$ is greater than 0.1.[16]

Most frequently[7-32] the magnitude of internal friction is determined from the damping of free vibrations of a sample, since for metals and alloys $\tan\varphi < 0.1$. To do this experimentally one obtains (if the frequency of vibrations exceeds 2 cps) a time display of the free oscillations of the sample (Fig. 1a), and from it one computes the logarithmic decrement of the damped vibrations θ (taken to be a measure of internal friction), which by definition is equal to

$$\theta = \ln \frac{a_n}{a_{n+1}}. \qquad (2)$$

Here a_n and a_{n+1} are the amplitudes of oscillations in the n-th and (n + 1)-th cycles.

Usually, to increase the accuracy of the determination of θ (particularly in the case of visual observations, which are possible only when the frequency of the free vibrations does not exceed 2 cps), one employs the formula

$$\theta = \frac{1}{N} \ln \frac{a_n}{a_m}, \qquad (3)$$

which gives the mean value of the logarithmic decrement for the portion of the time display bounded by the n-th and m-th cycles ($N = m - n$). If within this interval the value of the decrement varies slowly as the amplitude decreases, formula (3) introduces negligible errors. But if this formula is applied to the whole time display ($m = N$, $n = 0$), the calculated value of θ may differ appreciably from the mean values of the logarithmic decrement for a small portion of the time display, in view of its possible dependence on the amplitude of the oscillations.

Logarithmic decrements lower than 10^{-4} cannot be measured, since the losses in the apparatus, even in vacuo, are of the same order of magnitude. It is also impossible to measure the internal friction by this method when $\theta \geq 0.3$, for at these values of θ the motion of the sample becomes aperiodic.

The next method[34-38] of measuring internal friction consists of observing the amplitude of forced oscillations while the frequency of the external ex-

citing force is slowly varied and its amplitude is kept constant. The amplitude of the oscillation of the sample (Fig. 1b) reaches a maximum when the frequency of the external force is equal to the resonance frequency ν_0 of the sample, and falls off sharply if the frequency of the external force is either greater or less than this frequency. As a measure of internal friction in this method we adopt the quantity

$$B = \frac{\Delta\nu}{\nu_0}, \qquad (4)$$

where $\Delta\nu$ is the half-width of the resonance peak.

This method gives good results when the internal friction determined by the quantity B is not smaller than 10^{-5}, since it is difficult to obtain resonance curves when the internal friction is very small, owing to the great sharpness of resonance.

The internal friction can also be measured by other methods (cf. a review of these methods in references 3, 4, 49-52, 87, 250, and 251), but they are not often used.

The quantities indicated above as measures of internal friction are interrelated by simple equations, which are valid when the internal friction does not depend on the amplitude and when the magnitude of the internal friction, defined for example as $\tan\varphi$, does not exceed 0.1. Since in the preponderant majority of known cases the internal friction in metals and alloys does not exceed 0.1, these relations between δ, $\tan\varphi$, θ, and B do have a practical meaning: they enable us to compare results of investigations obtained by different methods. In references 3, 4, 5, and 51 the following relations between the measures of internal friction have been obtained:

$$\tan\varphi = \frac{\theta}{\pi} = \frac{\delta}{2\pi} = \frac{B}{\sqrt{3}}. \qquad (5)$$

When the internal friction is large, the relations between these different measures of internal friction depend on the mechanism of internal friction and must be derived separately for each case.[5]

We note that in the theory of oscillatory circuits the quantity $\sqrt{3}/B$ is usually referred to as the figure of merit and is denoted by Q. Therefore many authors use the symbol Q^{-1} for the magnitude of internal friction.

In conclusion we should say a few words about the earlier work (approximately prior to 1930) on internal friction. In these investigations[54-76] quite rough static and dynamic methods were used to find the coefficient of viscosity, which was used to characterize internal friction in materials. However, as Pines[77] has correctly noted, this should not be done, since the coefficient of viscosity and the internal friction are by no means the same. Therefore, unfortunately, it does not appear to be possible to make use of the results of most of the earlier research on internal friction.

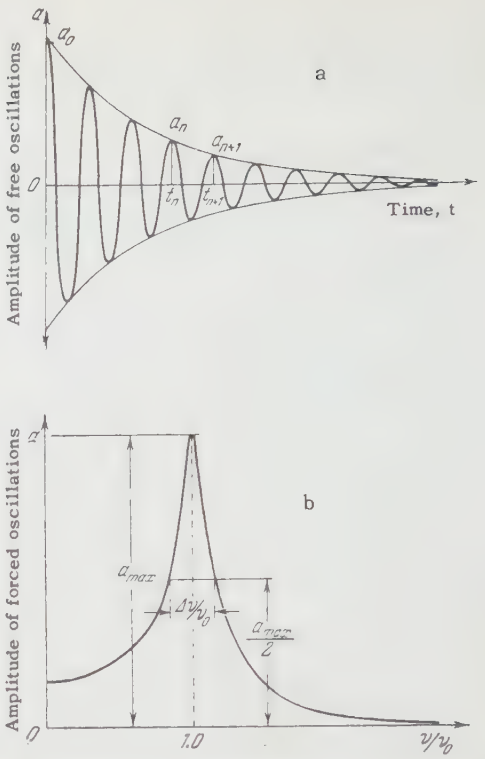


FIG. 1. Methods of determining internal friction: (a) by time display of free oscillations, (b) by width of the amplitude resonance peak.

Recently there has been a sharp increase in interest in the study of the dependence of internal friction on various factors, particularly with a view toward studying various problems of physical and chemical kinetics. Special attention has been devoted to the temperature dependence of internal friction, and the present review is devoted to this topic.

II. EXPERIMENTAL RESULTS

1. Pure Metals

Figures 2 and 3 show the temperature dependence, obtained by high frequency methods, of internal friction in tin,[79,252] lead,[80,81,82] aluminum,[81;83,249] silver,[80,81] and copper[80,81,84,85] at low temperatures. The maximum friction in tin is associated with the transition of tin into the superconducting state[251-253] at 3.73°K. A similar maximum occurs in the case of lead.[251] The peak of internal friction in the case of lead, aluminum (at approximately 105°K), and silver is explained largely by the motion of dislocations in the stress field. The peak for aluminum at approximately

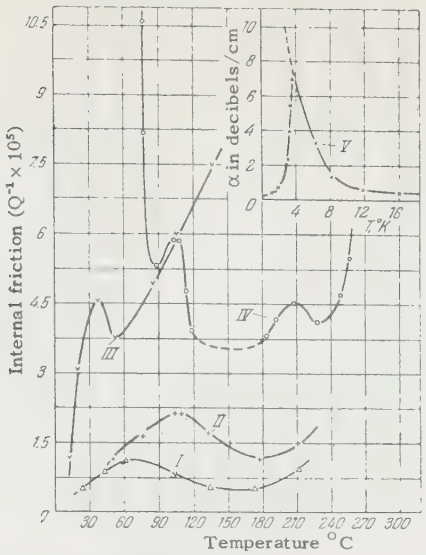


FIG. 2. Internal friction in polycrystalline lead (I – 99.9% Pb), aluminum (II, III – 99.99% Al), silver (IV – 99.5% Ag), and in a single crystal of tin (V – > 99.99% Sn) at low temperatures. Frequency of oscillations I, II, IV ~ 10^4 cps, III ~ 10^3 cps, V ~ 10^7 cps.

235° K is explained by the motion of single vacancies.[249] Near room temperature, the internal friction of all metals investigated so far increases monotonically. It is of interest to note the good agreement between the results of investigations by different authors[80,84] in the case of single crystals of copper, shown in Fig. 3 by the curves III′ and III″, which enables us to join the two branches into a single curve (dotted line in Fig. 3).

Figures 4 to 9 give the temperature dependence, obtained by various methods, of internal friction in tin,[86-87] lead,[88] zinc,[89] magnesium,[15,47] aluminum,[15,47,90,94] silver,[97,98] gold,[99,100] copper[93-97,121] nickel,[103,106] cobalt[103,104,107] iron,[103,104,108,109] titanium,[102-104] platinum,[98,101] zirconium,[112] molybdenum,[103,104] tantalum,[100,112] and tungsten[103,104] at

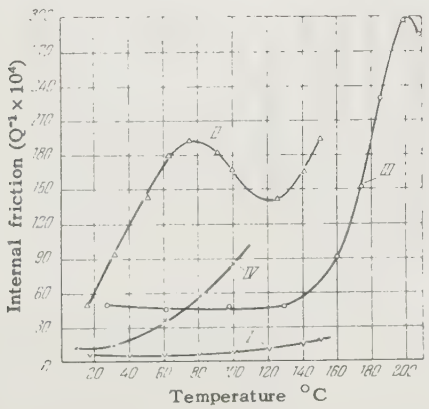


FIG. 4. Temperature dependence of internal friction in single-crystal and polycrystalline tin (I, II – 99.99% Sn, ν ~ 300 cps), polycrystalline lead (III – 99.9% Pb, ν ~ 950 cps), and zinc (IV – 99.6% Zn, ν ~ 180 cps).

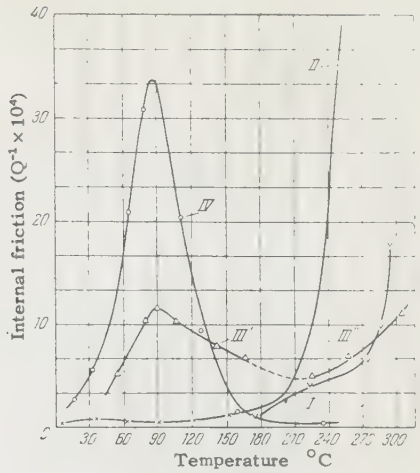


FIG. 3. Internal friction in single crystals of tin (I – 99.93% Sn, ν ~ 1.5×10^3 cps), copper (III′ – 99.9% Cu, ν ~ 4×10^4 cps, III″ – 99.86% Cu, ν ~ 3.9×10^4 cps) and in polycrystalline lead (II – 99.9% Pb, ν ~ 10^4 cps), copper (IV – 99.9% Cu, ν ~ 3×10^4 cps) at low temperatures.

temperatures above room temperature. The temperature dependence of internal friction of tin, lead, zinc, silver, and copper was measured at frequencies from 180 to 800 cps; in the case of aluminum, magnesium, and cobalt the frequency range was 1355 to 8×10^4 cps; in the case of magnesium, aluminum, gold, copper, nickel, cobalt, iron, titanium, platinum, zirconium, molybdenum, tantalum, and tungsten the frequency was on the order of 1 cps. All the metals were in polycrystalline form. Tin, aluminum, gold, copper, and iron were also used in single-crystal form. All the metals were annealed at high temperatures for a time long enough to remove internal stresses and to obtain a more or less equilibrium state.

It is seen from the figures that the internal fric-

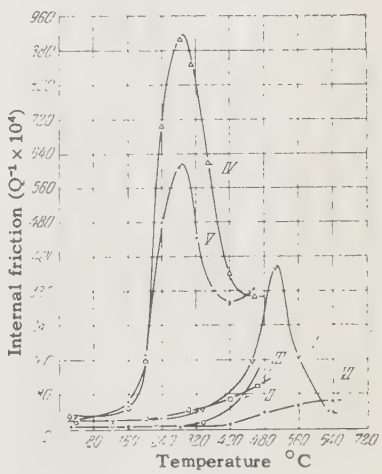


FIG. 5. Temperature dependence of internal friction in aluminum–single crystal (II – 99.99% Al, ν ~ 4×10^4 cps; III – 99.5% Al, ν ~ 1 cps) and polycrystalline (I – 99.99% Al, ν ~ 10^4 cps; IV – 99.991% Al, 1 cps; V – 99.98% Al, ν ~ 1 cps; VI – 99.99% Al, ν ~ 5×10^3 cps).

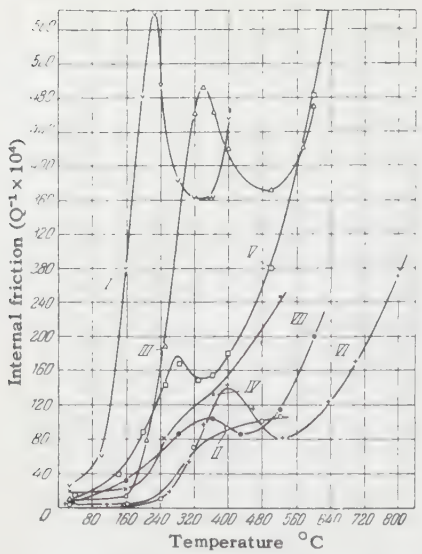


FIG. 6. Temperature dependence of internal friction in polycrystalline magnesium (I – 99.97% Mg, $\nu \sim 1$ cps; II ≈ 99.99% Mg, $\nu \sim 5 \times 10^3$ cps), copper (III – 99.999% Cu, $\nu \sim 1$ cps; IV – 99.999% Cu, very large crystals, $\nu \sim 1$ cps), and silver (V – 99.99% Ag, $\nu \sim 1$ cps).

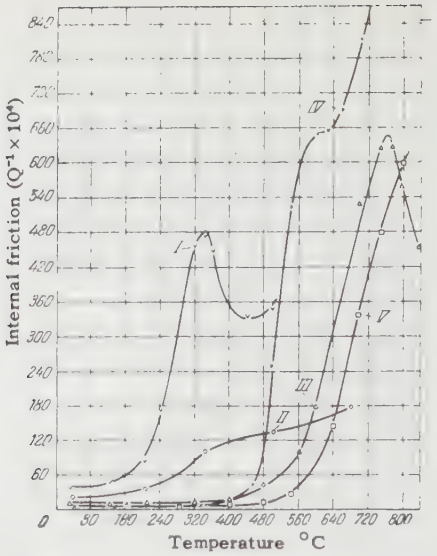


FIG. 7. Temperature dependence of internal friction in polycrystalline and single-crystal gold (I, II – 99.9998% Au, $\nu \sim 1$ cps), polycrystalline platinum (III – 99.97% Pt, $\nu \sim 1$ cps) and titanium (IV – 99.81% Ti, $\nu \sim 1$ cps; V – 99.81% Ti, very large grains, $\nu \sim 1$ cps).

tion of single crystals of metals increases monotonically with increasing temperature, without giving any maxima. On the other hand, the internal friction of polycrystalline metals (tin, lead, magnesium, aluminum, silver, gold, copper, nickel, iron, and zirconium) shows a maximum at temperatures (with the exception of tin and lead) equal to approximately (0.4 to 0.6) T_{melt}. Maxima in the internal friction of tin and lead are observed at temperatures ~(0.7 to 0.8) T_{melt}. A shift of the maximum towards comparatively higher temperatures is associated with an increase in the frequency of oscillations of the sample. The internal friction of titanium, molybdenum, tantalum, and tungsten does not have a maximum in the temperature region above 0.4 T_{melt}. For these not very pure metals the only characteristic feature is the existence of an inflection point in the curve $Q^{-1}(T)$. In Figs. 8 and 9 the inflection points for molybdenum, tantalum, and tungsten are not shown, since they lie above 800°C. The curve of $Q^{-1}(T)$ for tantalum in the low temperature region (~0.14 to

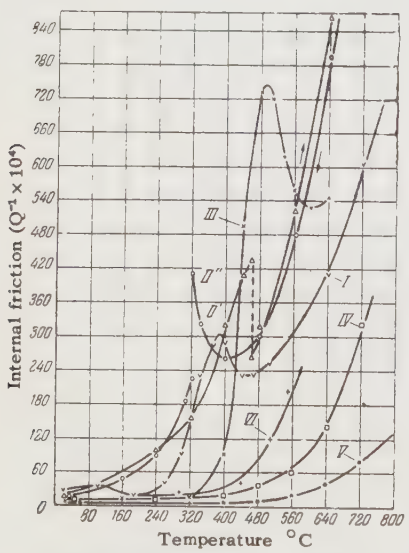


FIG. 8. Temperature dependence of internal friction in polycrystalline nickel (I – 99.99% Ni), cobalt (II′ – heating, II″ – cooling, 99.98% Co), iron (III, IV – 99.95% Fe, VI – very large grains), molybdenum (IV 99.92% Mo), and tungsten (V – 99.61% W). Frequency of oscillation $\nu \sim 1$ cps.

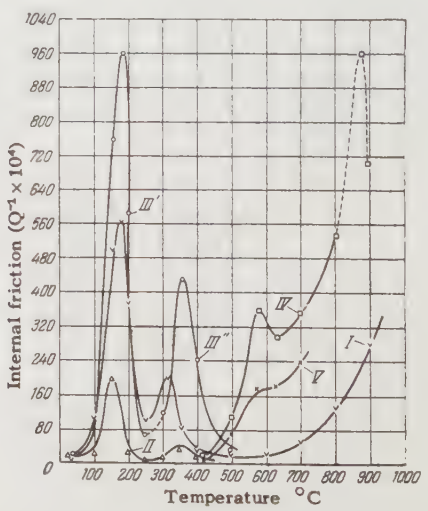


FIG. 9. Temperature dependence of internal friction in polycrystalline tantalum (I – 99.76% Ta + 0.17% Al, enriched by carbon, oxygen and nitrogen; II, III′, III″ – 99.9% Ta, enriched by carbon, oxygen and nitrogen) and zirconium (IV – 97.5% Zr + 2.4% Hf; V – the same sample, saturated with oxygen). Frequency of oscillation $\nu \sim 1$ cps.

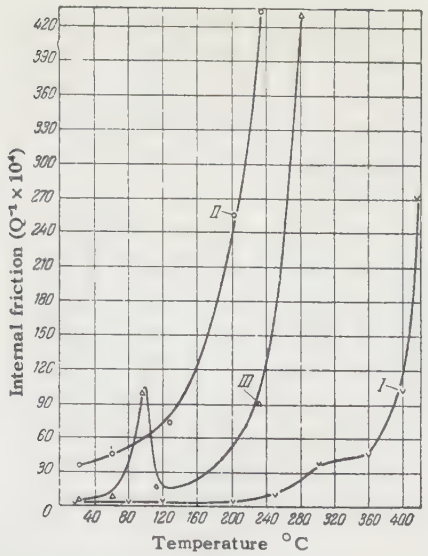


FIG. 10. Temperature dependence of solid solutions Cu – Bi (I – technical copper + 0.01% Bi, $\nu \sim 10^3$ cps), Pb – Sn (II – 99.9998% Pb + 0.058% Sn, $\nu \sim 3 \times 10^4$ cps), Pb – Bi (III – 99.9998% Pb + 0.053 Bi, $\nu \sim 3 \times 10^4$ cps).

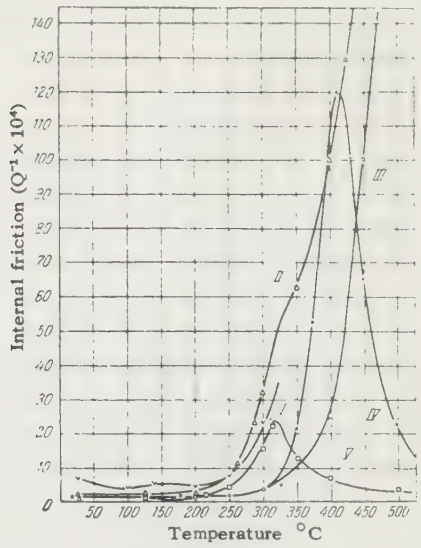


FIG. 11. Temperature dependence of solid solutions Al – Mg (I – 99.99% Al + 21.18% Ag, $\nu \sim 13 \times 10^3$ cps) and Cu – Zn (II – 99.9% Cu + 28% Zn, $\nu \sim 5 \times 10$ cps; III – Cu + 30% Zn, $\nu \sim 800$ cps; IV – Cu + 30% Zn, $\nu \sim 600$ cps, single crystal; V – Cu + 45% Zn, $\nu \sim 21 \times 10^3$ cps, single crystal).

0.19 T_{melt}) shows two sharply pronounced maxima, whose nature, as we shall see, differs from the nature of the maxima encountered at higher temperatures. Metals like cobalt and zirconium exhibit a very interesting dependence of internal friction on the temperature. For example, in the case of cobalt the internal friction rises rapidly with increasing temperature until it reaches a maximum of 450° C. In the neighborhood of this temperature (if the sample is maintained at constant temperature) the internal friction decreases rapidly and attains a certain stable value within approximately 30 minutes. As the temperature is increased further the internal friction continues to increase without showing any further peculiarities. On cooling, the internal friction decreases to 400° C, after which it begins to grow rapidly, reaching a maximum at 320° C. Isothermal conditioning (during 30 minutes) at this temperature leads to a rapid decrease in internal friction to a stable value. A further decrease in temperature leads to a smooth decrease in internal friction to $\sim 20 \times 10^{-4}$ at 20° C. In the temperature range from 300 to 450° a kind of "hysteresis" occurs in the temperature variation of internal friction.

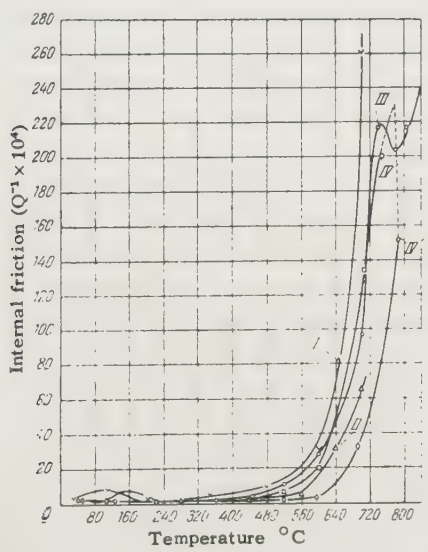


FIG. 12. Temperature dependence of Armco-iron (I) and of solid solutions Fe-Si (II – Armco-Fe + 4% Si), Fe – C (III – Armco–Fe + 0.84% C) and Fe–Co–Cr (IV′, IV″ Armco–Fe + 40% Co, 10% Cr; IV′ heating, IV″ cooling). Frequency of oscillation $\nu \sim 5 \times 10^3$ cps.

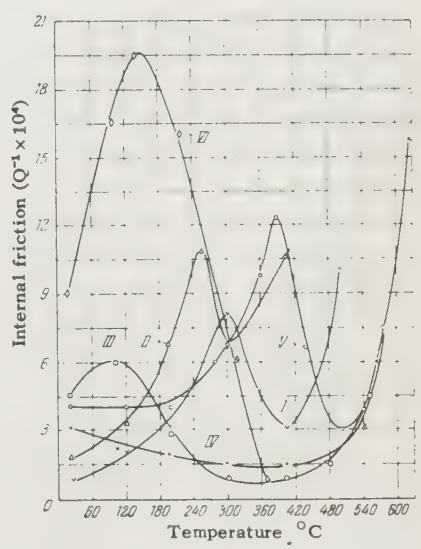


FIG. 13. Temperature dependence of internal friction in solid solutions Fe – Ni (I – Armco-Fe + 9.88% Ni + 0.67% Mn; II – Fe + 28.84% Ni + 0.38% Mn; III – Fe + 40% Ni; IV – Fe + 60% Ni, V – Fe + 80% Ni, VI – technical nickel). Frequency of oscillation $\nu \sim 10^3$ cps.

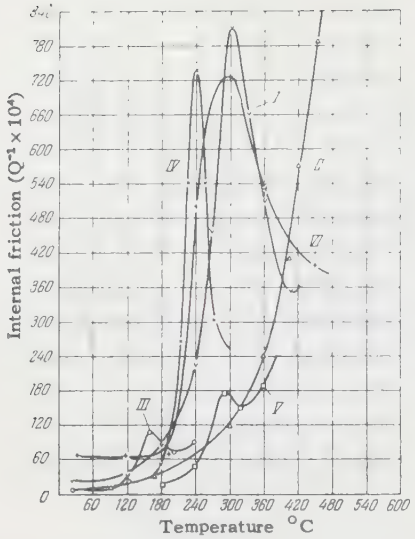


FIG. 14. Temperature dependence of internal friction in solid solutions Al – Cu (I – 99.991% Al + 0.5% Cu; II – Al + 4% Cu), Al – Ag (III – 99.99% Al + 20.1% Ag) and Ag – Zn (IV – 99.99% Al + 30.2% Zn; V – Al + 15.8% Zn). Frequency of oscillation $\nu \sim 1$ cps.

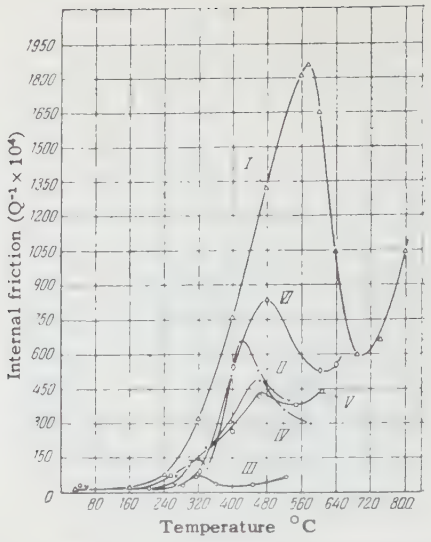


FIG. 15. Temperature dependence of internal friction in solid solutions Cu – Be (I – 99.992% Cu + 1.9% Be), Cu – Zn (II, III – cathode copper + 29.86% Zn; III – single crystal), Cu – Si (IV – 99.999% Cu + 0.1% Si), Cu – Al (V – 99.999% Cu + 0.1% Al) and Cu – Sn (VI – 99.9993% Cu + 6% Sn). Frequency of oscillation $\nu \sim 1$ cps.

Consequently, within a certain temperature range the internal friction in cobalt depends not only on the temperature, but also on the duration of the experiment. Zirconium behaves in a similar manner in the temperature range from 800 to 900° C. As is well known, these metals undergo transitions from one allotropic modification into another within these temperature ranges. The internal friction of titanium above 840° C should behave in a manner similar to that of cobalt and zirconium, but there are no experimental data.

The following conclusions can be drawn from the results of measurements shown in Figs. 4 to 9:

1. Well annealed metal samples exhibit temperature dependence of internal friction that is constant in time (reproducible in repeated measurements, if the oxidation of the samples is held down to a minimum).

2. The temperature dependence of internal friction of annealed single-crystal metal samples does not exhibit any maxima. The magnitude of the internal friction of single crystal samples is always less than the magnitude of the internal friction of corresponding polycrystalline samples.

3. Each polycrystalline metal (at a constant frequency of free vibrations) shows at least one maximum or inflection point on the curve of the temperature dependence of internal friction, provided no allotropic transitions take place on heating.

4. The internal friction of metals which may undergo allotropic changes on heating (for example, cobalt or zirconium), is not constant in time in the temperature range in which an allotropic change occurs. An "hysteresis" of internal friction is observed within this temperature range.

5. The maximum (or inflection point) of internal friction lies for the most part in the temperature range of recrystallization of the metal. Its position for a given metal, as shown by experiment, depends on the magnitude of the preceding plastic deformation: an increase in the degree of plastic deformation shifts the maximum towards lower temperatures. For different metals subjected to the same degree of plastic deformation, the position of the maximum (or inflection point) of internal friction is determined to some extent by the melting point of the metal: for metals of higher melting points the maximum of internal friction occurs at higher temperatures. The position of the maximum is also influenced by the frequency of oscillations: as the frequency of oscillations of the sample is increased the maximum is observed at increasingly higher temperatures.

6. At a given temperature metals of higher melting point exhibit a lower value of internal friction compared to that for metals of lower melting point. This difference increases as the temperature increases.

7. The magnitude of internal friction depends also on the preceding deformation. As a rule, the internal friction is considerably greater in the plastically deformed metals than well annealed ones.

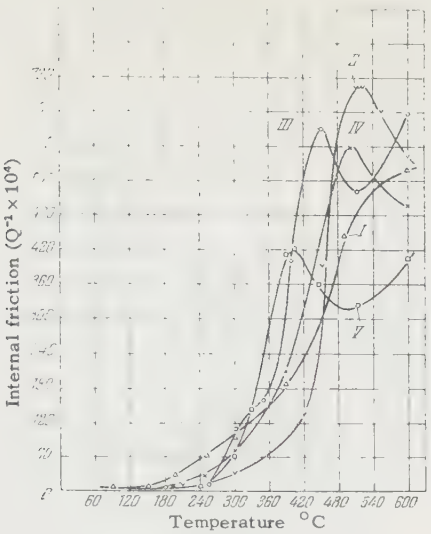


FIG. 16. Temperature dependence of internal friction in solid solutions Cu – Zn (I – spectral purity Cu + 1.17% Zn; II – 8.01% Zn; III – 32.1% Zn) and Cu – Ga (IV – 0.97.% Ga; V – 16.2% Ga). Frequency of oscillation $\nu \sim 1$ cps.

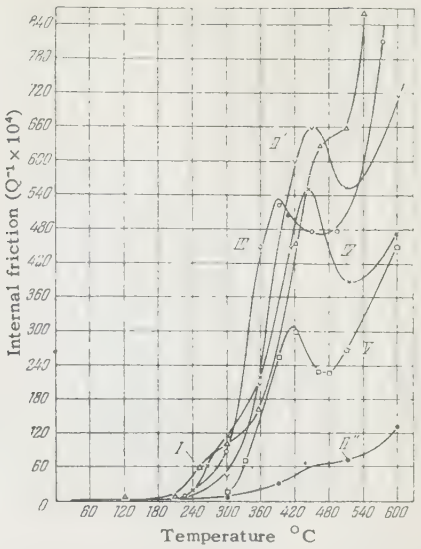


FIG. 17. Temperature dependence of internal friction in solid solutions Cu – Ge (I – spectral purity Cu + 1.02% Ge; II′, II″ – 4.94% Ge; II – very large grains), Cu – As (III – 3.93% As) and Cu – Si (IV – 0.9% Si; V – 9.0% Si). Frequency of oscillation $\nu \sim 1$ cps.

2. Alloys

Figures 10 to 13 present results of investigations of the temperature dependence of internal friction of binary alloys Pb-Sn,[82] Pb-Bi,[82] Al-Ag,[47] Cu-Bi,[43] Cu-Zn,[114-117] Fe-Si,[118] Fe-C,[11,118,119] and Fe-Ni,[120-121] and of the ternary alloy Fe-Co-Cr[121] at temperatures above room temperature, obtained by different methods at frequencies from 600 to 3×10^4 cps.

Figures 14 to 23 present the temperature dependence of the internal friction of the following binary alloys: Al-Zn,[123] Al-Mg,[124,125] Al-Ag,[126,127] Al-Cu,[128,129] Cu-Ga,[122] Cu-Sn,[95] Cu-Zn,[122,130,131] Cu-Al,[132] Cu-As,[122] Cu-Ge,[122] Cu-Si,[122,132] Cu-Be,[133-135] Au-Ni,[136] Ni-C,[103,105,106] Ni-B,[137] Ni-Be,[137] Ni-Mn,[137] Ni-Fe,[137] Ni-Ti,[137] Ni-Cr,[137,168] Ni-Zr,[137] Ni-Nb,[137] Ni-Mo,[137] Ni-W,[137] Fe-B,[138] and Fe-W,[146] and of the complex alloys: kovar,[137,139] nichrome,[137] alloy No. 2,[137] elinvar,[137] elgiloy,[137] Nimo alloy,[137,139] nimonik-80,[137] alloy 38KhMYuA,[140] Fe-Gr-Ni alloy,[141-144] and austenite steel 25-20.[145] The measurements have been carried out over a wide range of temperatures by various low-fre-

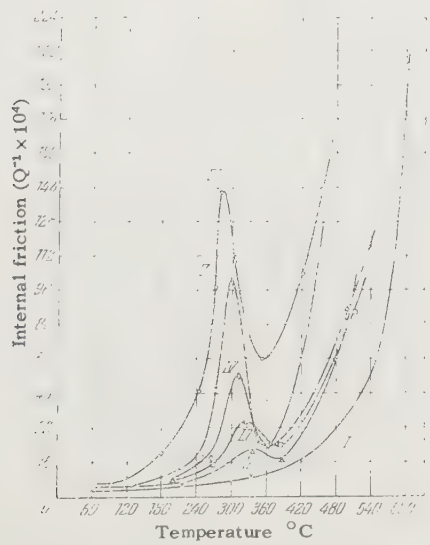


FIG. 18. Temperature dependence of internal friction in solid solutions Cu – Zn (I – spectrally pure Cu + 5% Zn; II – 10%, III – 15%, IV – 20%, V – 25%, VI – 30% Zn; I and II very large grains). Frequency of oscillation $\nu \sim 1$ cps.

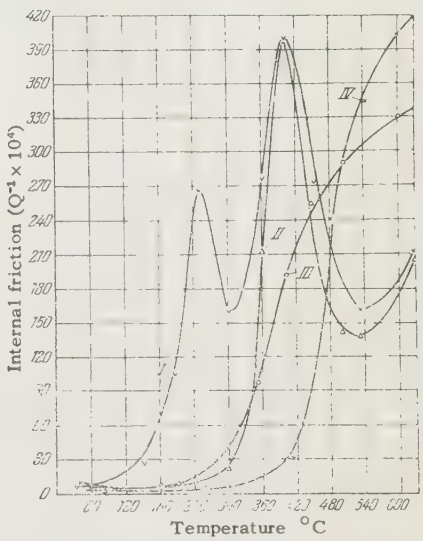


FIG. 19. Temperature dependence of internal friction in solid solutions Au – Ni (I – 99.98% Au + 30% Ni, sample is quenched in oil at 800°C; II – the sample is heated at 400° and is rapidly cooled in vacuo; III – the sample is kept at 400° for 240 hours, and is then slowly cooled), and Ni – C (IV – 99.99% Ni + 0.21% C). The frequency of oscillation is $\nu \sim 1$ cps.

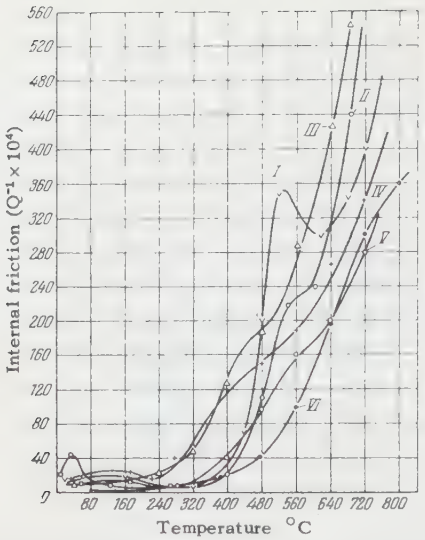


FIG. 20. Temperature dependence of internal friction in solid solutions Ni – B (I – 99.99% Ni + 0.25% B), Fe – B (II – Armco-Fe + 0.006% B), Ni – Fe (III – 99.99% Ni + 2.6% Fe), Ni – Zr (IV – 99.99% Ni + 5.5% Zr) and Ni – Nb (V – 99.99% Ni + 1.76% Nb. Frequency of oscillation $\nu \sim 1$ cps.

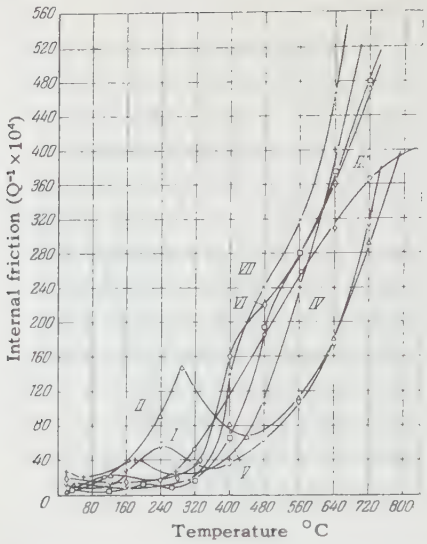


FIG. 21. Temperature dependence of internal friction in solid solutions Ni – Be (I, II – 2.9% Be; I – slowly cooled, II – quenched in water from 950°C), Ni – Ti (III – 1.04% Ti), Ni – Mn (IV – 22.8% Mn), Ni – Cr (V – 1.19% Cr), Ni – Mo (VI – 0.86% Mo) and Ni – W (VII – 1.54% W) based on 99.99% nickel. Frequency of oscillation $\nu \sim 1$ cps.

quency methods. All the alloys, with a few exceptions (cf. Figs. 14, 15, and 17), were investigated in their polycrystalline state.

We also note that within a narrow temperature range (from 20 to 300 or 400°) investigations have been carried out for the pure metals zinc,[29,49] silver,[149] copper,[147,149] and nickel,[35] and also for the following alloys: Sb-Bi,[150] Al-Cu,[151] Al-Cu-Si,[151] Al-Cu-Mg,[151] Al-Mg-Si,[151] Al-Cu-Mg-Si,[151] Ag-Cd,[186] Ag-Zn,[186] Au-Cd,[187] Au-Zn,[186] Au-Cu,[185] Cu-Zn,[185,186] Mn-Cu,[45] Fe-B,[152] Fe-C,[153-164,217,247,248] Fe-N,[153,155-159] Fe-Ni,[165] 105-Kh-12 steel,[166] Pd-H,[167] Pd-Cu,[185] and Pt-Cu.[185] The investigations were carried out by different methods, primarily in order to find out the behavior of the element dissolved in the alloy.

The measurements of the temperature dependence of internal friction of the pure metals and alloys enumerated above lead to the following conclusions:

1. The internal friction of alloys increases with increasing temperature (over a wide range

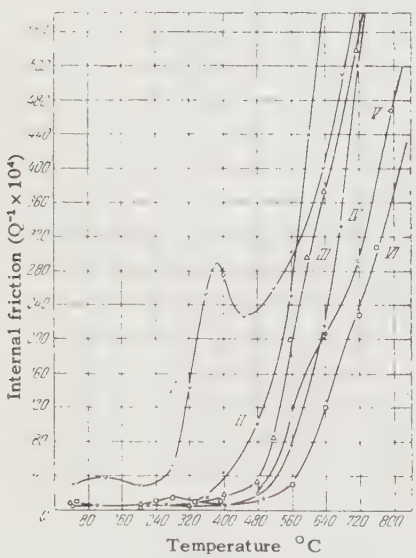


FIG. 22. Temperature dependence of internal friction in pure nickel (I – 99.99% Ni), kovar (II), technical nichrome (III), alloy No. 2 (IV), elinvar (V), elgiloy (VI – alloy K = 40). Frequency of oscillation $\nu \sim 1$ cps.

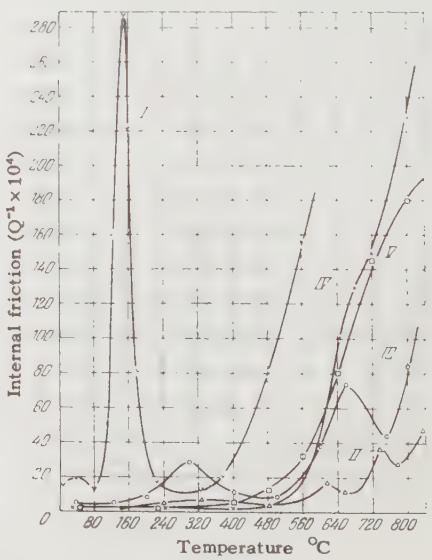


FIG. 23. Temperature dependence of internal friction in alloy ZVKhMYuA (I), austenite steel 18-18 (II), austenite steel 25-20 (III – containing 0.3% C, sample quenched in water from 1230°C), Nimo alloy (IV), and alloy nimonik-80 (V). Frequency of oscillation $\nu \sim 1$ cps.

of temperatures) much more slowly than the internal friction of those pure metals of which the alloys are formed.

2. The magnitude of internal friction, particularly at high temperatures, is less for the alloys than for the pure metals.

3. At high temperatures the magnitude of internal friction turns out to be related to the stability of the alloy at the same temperatures. In all the investigations, the magnitude of internal friction was found to be lower for the less refractory alloys.

4. All the binary alloys investigated exhibit at a certain temperature one maximum or inflection point in the curve of $Q^{-1}(T)$. More complex alloys may have several such points.

5. Just as in the case of pure metals, the maxima of internal friction in the case of an alloy are displaced towards higher temperatures (or else disappear entirely) as the frequency of oscillations of the sample is increased.

6. As a rule, at higher temperatures deformed alloys have a greater internal friction the same samples when previously well annealed.

III. THE THEORY OF INTERNAL FRICTION

Any atomic system taken out of the state of equilibrium and left to itself must either return to the initial state or go over into a new equilibrium state, if several stable states correspond to the given conditions. The transition of a system from one equilibrium state to another is characterized, as is well known, by a relaxation time (one or several). If a periodic force is applied to the system, there will be observed a dissipation of elastic energy which depends on the frequency of the applied force. This dissipation of elastic energy of the oscillations is determined by the totality of all the relaxation phenomena called into play when the atomic system is subjected to a periodic deformation, and therefore cannot as yet be determined in general form. In the first approximation (when the relaxation phenomena occur independently of one another) we can assume that the total internal friction, which characterizes the dissipation of the energy of the oscillations, is determined by the sum of the contributions made by the different relaxation phenomena.[78,93,137]

Let us consider the basic relaxation phenomena which occur in a solid when it is deformed.

1. The General Thermodynamic Theory of Internal Friction

We consider a homogeneous isotropic solid. For small deformations, for small temperature changes, and for small deviations from the equilibrium state the stress tensor σ_{ik} is determined by the following equation:[169-171]

$$\sigma_{ik} = K\varepsilon_{ll}\delta_{ik} + 2\mu\left(\varepsilon_{ik} - \frac{1}{3}\varepsilon_{ll}\delta_{ik}\right) - \alpha K(T - T_0)\delta_{ik} + \frac{2\eta_1}{\tau_1}\int_{-\infty}^{t} e^{\frac{(t'-t)}{\tau_1}}\left[\dot{\varepsilon}_{ik}(t') - \frac{1}{3}\dot{\varepsilon}_{ll}\delta_{ik}\right]dt' + \frac{\eta_2}{\tau_2}\int_{-\infty}^{t} e^{\frac{(t'-t)}{\tau_2}}\dot{\varepsilon}_{ll}(t')\delta_{ik}\,dt'. \quad (6)$$

In the case of isothermal pure shear (for example, small torsional oscillations of a rod undergoing relaxation) $\epsilon_{ll} = 0$, $\Delta T = 0$, and therefore

$$\sigma_{ik} = 2\mu\varepsilon_{ik} + \frac{2\eta_1}{\tau_1}\int_{-\infty}^{t} e^{\frac{t'-t}{\tau_1}}\dot{\varepsilon}_{ik}(t')\,dt'. \quad (7)$$

If the deformation is a periodic function of the time,

$$\varepsilon_{ik} = \varepsilon_{ik}^{(0)} e^{i\omega t} \quad (8)$$

and the generalized Hooke's law can be written

$$\sigma_{ik} = \mu^* \varepsilon_{ik}, \quad (9)$$

where the complex shear modulus is given by

$$\mu^* = \mu + \frac{i\omega\eta_1}{1 + i\omega\tau_1}. \quad (10)$$

On multiplying the numerator and denominator of the second term in the right hand side of (10) by $(1 - i\omega\tau_1)$ we shall obtain

$$\mu^* = \mu + \frac{(\omega\tau_1)^2}{1+\omega^2\tau_1^2}\frac{\eta_1}{\tau_1} + \frac{i\omega\tau_1}{1+\omega^2\tau_1^2}\frac{\eta_1}{\tau_1}. \quad (11)$$

For the dynamic modulus one usually takes the real part of the complex modulus, i.e.,

$$\mu_\omega = \mu + \frac{(\omega\tau_1)^2}{1+\omega^2\tau_1^2}\frac{\eta_1}{\tau_1}. \quad (12)$$

From the last expression we obtain

$$\left.\begin{aligned}\mu_\infty &= \mu + \frac{\eta_1}{\tau_1},\\ \mu_0 &= \mu.\end{aligned}\right\} \quad (13)$$

Taking (12) and (13) into account, we obtain the shear-modulus decrement that defines the degree of relaxation:

$$\Delta_\omega = \frac{\mu_\infty - \mu_\omega}{\mu_0} = \frac{\Delta_0}{1+\omega^2\tau_1^2}, \quad (14)$$

where

$$\Delta_0 = \frac{\mu_\infty - \mu_0}{\mu_0}$$

is the maximum value in the decrement of the shear modulus. For the measure of internal friction one usually takes[3,4,51,52,78] $\tan\varphi$, which is equal to the

ratio of the imaginary part of the complex modulus to its real part. By taking this ratio from (11) we obtain

$$Q^{-1}(\omega) = \frac{\Delta_0 \omega\tau_1}{1 + \frac{\mu_\infty}{\mu_0}\omega^2\tau_1^2}. \qquad (15)$$

Expressions (14) and (15) agree [up to a factor μ_∞/μ_0 in the denominator of (15)] with the expressions obtained by Zener,[3] where μ_∞ denotes the "unrelaxed" while μ_0 denotes the "relaxed" modulus. A similar result was recently obtained by V. T. Shmatov.[234]

Figure 24 shows the variation in the internal friction $Q^{-1}(\omega\tau)$ and of the ratio of the dynamic modulus μ_ω to the unrelaxed modulus μ_∞ as functions of the parameter $\omega\tau$ in the case when $\mu_0/\mu_\infty = 0.6$. It can be seen that the magnitude of the internal friction attains a maximum (whose value depends on Δ_0 and on the ratio μ_∞/μ_0 for the given material) at $\omega\tau = \sqrt{\mu_0/\mu_\infty} \approx 1$, i.e., when the frequency ω of the oscillations of the sample approximately coincides with the reciprocal value of the relaxation time τ for the process. The ratio μ_ω/μ_∞ at the point where $Q^{-1}(\omega\tau)$ is a maximum has the greatest rate of variation with respect to the parameter $\omega\tau$.

For $\omega\tau > 10^{-1}$ the internal friction is insignificant, and the variation of the modulus is practically equal to zero, i.e., the dynamic modulus is practically equal to the "unrelaxed" modulus μ_∞. Consequently, in this frequency range $[\omega > (10\tau)^{-1}]$ there is practically no relaxation, and therefore there is no appreciable dissipation of the elastic energy of oscillations. For $\omega\tau < 10^{-1}$, the dynamic modulus is practically equal to the "relaxed" modulus μ_0, while the internal friction is again very small. In the intermediate frequency range we have a partial relaxation of the dynamic modulus and a considerable internal friction, which reaches a maximum at $\omega\tau \cong 1$.

Thus, according to the thermodynamic theory, in the case of a homogeneous isotropic solid, an appreciable dissipation of elastic energy accompanying the periodic deformation occurs whenever the reciprocal value of the relaxation time of the relaxation process under consideration coincides in order of magnitude with the frequency of oscillation of the samples. Phenomena, with one or several relaxation times may be atomic diffusion, relaxation of magnetic flux, heat conductivity, interaction of phonons with the conduction electrons, etc. Thus, by varying the frequency of oscillation of the sample from very low values ($\sim 10^{-10}$ cps in the case of atomic diffusion) up to very high ones ($\sim 10^{10}$ cps in the case of interaction of phonons with conduction electrons) one can obtain a discrete relaxation spectrum even for a homogeneous isotropic solid.

Real solids, as a rule, are neither homogeneous nor isotropic. Therefore in the case of each of these phenomena we may find for these solids not one relaxation time, but a whole set, and in the general case of a continuous spectrum. If several relaxation processes are taking place in the solid simultaneously, each of which can be characterized by its own particular relaxation time τ_i, then the internal friction in such a solid is defined by the expression[3,4,78,93,103,104,137]

$$Q^{-1}(\omega) = \sum_{i=1}^{n} \frac{\Delta_i^0 \omega\tau_i}{1 + \left(\frac{\mu_\infty}{\mu_0}\right)_i (\omega\tau_i)^2}. \qquad (16)$$

The frequency dependence of the internal friction of such a solid may have not a single maximum, but several, or, if the relaxation times τ_i are grouped around some one mean value τ_{av}, then it may have a single maximum which is considerably "smeared."[93,103,104,137] Such "smeared" maxima can be observed in the case of a curve of the temperature dependence of internal friction of polycrystalline metals. To explain such maxima, the idea of a "two component system" was introduced.[3,180,181,193] This idea consists of assuming that the solid comprises two phases, one amorphous and the other perfectly elastic. In such a solid the stress relaxes with a relaxation time

$$\tau = \frac{\eta}{\mu}, \qquad (17)$$

where η is the coefficient of viscosity of the amorphous phase, and μ is the shear modulus of

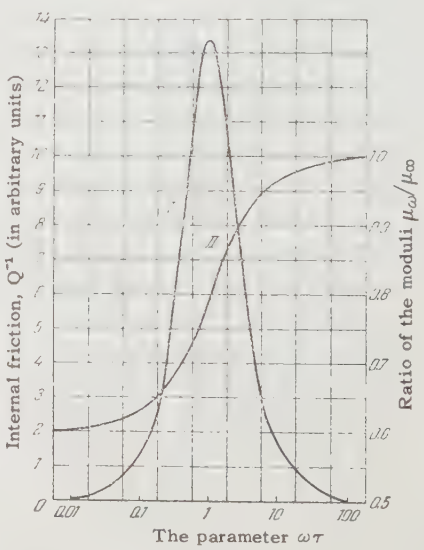


FIG. 24. Dependence of internal friction (I) and of ratio μ_ω/μ_∞ (II) on the parameter $\omega\tau$.

the elastic phase. This relaxation time will depend on the physical nature of the amorphous region, and also on its dimensions and shape. It may be easily shown that if the amorphous region has a width c and a length l, the coefficient of viscosity of this region is related to the relaxation time by the following equation[15]

$$\eta_1 = \frac{\mu_\infty \tau}{l} c, \tag{18}$$

from which it can be seen that the relaxation time depends on the dimensions of the amorphous region (and on its shape).

If amorphous regions are embedded into an elastic matrix which is the base of the solid, then each such region will relax the stress according to its own relaxation time τ_i. The behavior of such a solid will thus be determined by the totality of all the relaxation times τ_i. This is what causes the "smearing" of the maxima.

Experiments show that the role of amorphous regions can be played by slip bands or by grain boundaries.

Finally, in the case of a continuous distribution of relaxation constants τ, the internal friction is given by the expression[3,37]

$$Q^{-1}(\omega) = \int_{-\infty}^{+\infty} \frac{\psi(\tau)\,\omega\tau}{1 + \left(\frac{\mu_\infty}{\mu_0}\right)_{\text{av}} (\omega\tau)^2}\, d(\ln\tau), \tag{19}$$

where $\Psi(\tau)$ is an unknown distribution function, and difficult to determine experimentally. It can be found from the condition

$$(\Delta_0)_{\text{av}} = \int_{-\infty}^{+\infty} \psi(\tau)\, d(\ln\tau), \tag{20}$$

where $(\Delta_0)_{\text{av}}$ is the experimentally-known mean decrement in the modulus; this requires the inversion of the integral, which introduces a considerable inaccuracy into the final result.

In the case of relaxation phenomena associated with the diffusion displacement of atoms in the stress field, the relaxation time depends on the temperature, as is shown by numerous experiments, in accordance with the following law

$$\tau = \tau_0 e^{H/kT}, \tag{21}$$

where H is the heat of activation of the process under investigation. Since the internal friction Q^{-1} depends on the frequency and on the relaxation time through the parameter $\omega\tau$, we may leave the frequency ω unchanged, and vary the relaxation time τ by varying the temperature: at a certain temperature T we obtain the equality $\omega\tau \cong 1$ at which we observe the maximum internal friction due to the existence of the relaxation phenomenon under consideration. Thus it becomes possible to understand why at a certain temperature (for a given frequency ω) a maximum appears on the curve of the temperature dependence of internal friction. It also becomes possible to understand the shift of the maximum as the frequency of oscillation of the sample is increased.

The heat of activation H is determined in the following manner. Suppose that we have two curves of the dependence of internal friction on the temperature, one obtained at a constant frequency ν_1 and the other at a constant frequency ν_2. If we select on these two curves two points corresponding to the maxima of internal friction, this will mean that

$$\nu_1\tau_1 = \nu_2\tau_2$$

or, according to relation (21),

$$\nu_1 e^{H_1/RT_1} = \nu_2 e^{H_2/RT_2}.$$

By taking logarithms of both sides of the above equation we obtain, after a simple rearrangement,

$$H = R\frac{T_1T_2}{T_2 - T_1}\ln\frac{\nu_2}{\nu_1}. \tag{22}$$

As we shall see below, everything said with respect to the determination of H applies also to the case when the relaxation process is characterized not by a single relaxation time, but by a whole set of relaxation times grouped around some mean value τ_{av}. In this case we obtain H_{av} for the complex process under consideration. The quantity Δ_0 is determined experimentally[235,236] or is calculated from expression (15) in terms of known values of τ, μ_∞/μ_0, ω, and $Q^{-1}_{\max}$. In certain cases, as we shall see below, τ and Δ_0 can be estimated theoretically.

We now consider individual processes which give rise to internal friction, and start with processes whose relaxation times at room temperature are very large (of the order of 10^{10} sec and greater).

2. Internal Friction Due to the Ordering of Atoms under Stress

(a) **Substitution Alloys.** In annealed substitution solutions not undergoing ordering, the dissolved atoms are distributed isotropically in the lattice of the solvent. If a stress is applied to such an alloy, a certain amount of ordering in the distribution of the dissolved atoms is introduced.[4] The relaxation time τ associated with ordering under stress must be simply related to the mean frequency of atomic

transitions, since the mechanism of ordering is the same as in the case of volume diffusion. Therefore the relaxation time of the ordering process may be estimated by means of the following formula[237]

$$\tau = \tau_0 e^{H/RT}, \tag{23}$$

where H is the activation energy for diffusion. A noticeable maximum of internal friction in such solutions, due to a change of order in the distribution of the dissolved atoms under the influence of an alternating stress, can be observed only at comparatively high concentrations (of the order of 10 atomic per cent and higher) of the dissolved component and at comparatively high temperatures, when the reciprocal of the relaxation time τ^{-1} is comparable with the frequency of variation of the applied stress. Ordinarily low-frequency methods ($\nu \sim 1$ cycle/sec) are used in such cases. The internal friction in substitution alloys due to ordering under stress was first observed by Zener[116] in the case of single crystals of α-brass at ~600 cps. Later Kê and Nowick[131,182,238] confirmed the existence of such a peak in α-brass and in the Ag-Zn alloy at lower frequencies. In a detailed paper[238] Nowick studied the dependence of the internal friction of the Ag-Zn alloys on composition and on temperature.

Zener[237] gave an explanation of the above effect, based on the violation of cubic symmetry by a pair of different atoms forming the solid solution. Under the action of the applied stress, these pairs of atoms take up a preferred orientation. The tendency of a lattice containing such pairs to go over to a new equilibrium state is what gives rise to the appearance of internal friction when it is periodically deformed. The degree of relaxation in this case can be determined by means of the formula[237]

$$\Delta_0 = nE_T \left(\frac{\partial \varepsilon}{\partial n_p}\right)^2 \frac{1}{4kT}, \tag{24}$$

where n is the number of atom pairs per unit volume, E_T is the "isothermal" modulus of elasticity, ϵ is the tension strain, and

$$n_p = n \frac{e^{-u/kT}}{1 + e^{u/kT}}$$

is the number of pairs oriented along the new direction of p under the action of the applied stress.

If in the absence of stresses a certain degree of long range order exists in the solid solution, then the applied stress will give rise to a change in the degree of ordering. Relaxation processes of such type were first considered by V. S. Gorskii[183,184] in the case of the Cu-Au alloy.

The degree of relaxation for alloys undergoing ordering may be approximately defined[3] by

$$\Delta_0 = E_S \left(\frac{\partial S_c}{\partial T_c}\right)_\sigma \left(\frac{\partial \varepsilon}{\partial S_c}\right)_\sigma^2, \tag{25}$$

where E_S is the "adiabatic" modulus of elasticity, S_c is the entropy of displacement, T_c is the temperature equal to the derivative $(\partial H/\partial S_c)_\sigma$ at constant stress σ, H is the heat content per unit volume, and ϵ is the tension strain. The relaxation time is estimated by means of formula (23). In constrast to the case discussed above in which the solid solution is not ordered in the absence of stresses, we can obtain here a characteristic peak of internal friction considered only as a function of frequency (at T = const), but not as a function of temperature (for ω = const), because the degree of ordering in the absence of stresses is itself a function of the temperature. It is therefore difficult to explain the results of experiments in which a dependence of internal friction on the temperature has been obtained.[94,185,187] In these experiments, carried out using alloys undergoing ordering, it was shown that the internal friction increases appreciably near the Curie point for long-range order and continues to grow as we go to the temperature of disordering, reaching very large values. The explanation of relaxation phenomena in substitution alloys in terms of the reorientation of pairs of dissolved atoms has met objections on the part of Nowick[4] who, however, did not advance any serious arguments against the admissibility of such an explanation. In essence all his remarks reduced to certain improvements of the relaxation mechanism proposed by Zener. According to Nowick[4] the relaxation time is equal to

$$\tau = \alpha \Gamma_A^{-1}, \tag{26}$$

where Γ_A is the mean frequency of atomic transitions of the less mobile atoms of type A in the solution A-B, and α is a coefficient of proportionality. The value of Γ_A depends on the temperature according to the formula

$$\Gamma_A = \Gamma_0 e^{-H/RT}, \tag{27}$$

where H is the activation energy of volume diffusion. This energe depends strongly on the concentration of atoms of type A in the solution.[4,123,128] The degree of relaxation Δ_0 was not determined by Nowick, but it can be estimated approximately by using formula (30).

(b) Interstitial Solutions. In an interstitial solid solution free of stress the distribution of dissolved atoms is isotropic with respect to each atom of the

basic lattice. This means that the nearest dissolved atom has the same probability of being situated along any one of the possible crystallographic directions. For example, in a body-centered cubic lattice the interstitial atoms are mostly found not at points with the greatest free volume, but in positions of the type ($\frac{1}{2}$, 0, 0) (i.e., in the middle of the edges of the cell) or in the centers of the faces ($\frac{1}{2}$, $\frac{1}{2}$, 0) which are equivalent to the former positions. The interstitial atoms situated at such points deform the lattice, giving rise to deformations having tetragonal symmetry. The greatest deformations occur in the direction between two nearest atoms of the solvent, i.e., along one of the principal axes (1, 0, 0). In the absence of stresses all three types of interstitial positions are occupied in a disordered way, and therefore each type of these interstitial positions is associated with one-third the total number of interstitial atoms. If tension stress is applied, for example, along the z axis, the isotropic distribution of atoms will be violated; the equilibrium distribution now will be such that the greater number of the dissolved atoms is found in positions with the tetragonal axis z.

The transition to the new equilibrium situation is characterized by a certain relaxation time τ comparable with the lifetime of the interstitial atom in the interstitial position:[4,181]

$$\tau \cong \frac{1}{\nu_0} e^{H/RT}, \tag{28}$$

where ν_0 is the frequency of oscillation of the interstitial atom in the interstitial position, H is the heat of activation of a mol of interstitial atoms, which depends on the type of interstitial position and on the nature of both the interstitial atoms and the atoms of the solvent.

The following expression has been proposed[4] for the degree of relaxation of this process

$$\Delta_0 = \frac{2}{9} E_\infty \frac{\lambda^2 n}{kT}. \tag{29}$$

Here $\lambda = \partial\epsilon/\partial p$; p is an internal parameter, equal in our case to $n_z - n/3$, n is the total number of interstitial atoms per unit volume, n_z is the number of atoms in position z, k is the Boltzmann constant, E_∞ is the unrelaxed modulus of elasticity, and T is the absolute temperature. The internal friction due to the rearrangement of interstitial atoms under stress was first observed[239] and explained[159] in the case of α-iron. Later the same phenomenon was observed also in the case of other solid interstitial solutions having a cubic body centered lattice similar to solutions in α-iron. Thus, Kê investigated relaxation effects associated with the presence of carbon, oxygen,[111] and nitrogen[112] in tantalum. The behavior of carbon in niobium has been investigated by Wert.[240] Comparatively recently, papers have appeared on the study of relaxation phenomena associated with the behavior of interstitial atoms in face centered cubic lattices.[105,106,141-144,241]

In considering relaxation phenomena in solid solutions we have assumed that the dissolved atoms form a homogeneous solid solution at the temperature at which the investigations were made. However, certain metals that are completely soluble in each other at high temperatures no longer mix below a certain critical temperature. In such a case the degree of relaxation, which is defined in the general case by means of[3,83,184]

$$\Delta_0 = E_\infty \left(\frac{\partial \varepsilon}{\partial c}\right)_\sigma^2 : \left(\frac{\partial^2 F}{\partial c^2}\right)_\sigma, \tag{30}$$

(E_∞ is the unrelaxed modulus of elasticity, c is the atomic concentration of the dissolved component, and F is the free energy per unit volume) may be either anomalously small or very large, depending on the extent of the region of solubility, and on the temperature at which the measurement is made.

The curvature $(\partial^2 F/\partial c^2)_\sigma$ is always anomalously large in those phases which have a narrow range of solubility. Consequently, in such phases the relaxation effects associated with atomic diffusion under stress are anomalously small. Conversely, in the case of a wide range of solubility the curvature at a certain value of concentration, particularly near the critical temperature, is close to zero, and the degree of relaxation is anomalously large.

3. Internal Friction Due to Ferromagnetism

Because of the well-known connection between magnetic and mechanical phenomena, all types of magnetic relaxation lead to additional appreciable dissipation of elastic energy of oscillation. It is customary[175-178] to divide "magnetic" losses into:

(1) losses due to macroscopic eddy currents;

(2) losses due to microscopic eddy currents;

(3) losses associated with magnetomechanical hysteresis.

Losses of the first type are due to the relaxation of magnetic flux caused by the appearance of eddy currents produced by the change in the induction in the sample as a whole as a result of its periodic deformation. The relaxation time is determined in this case by the following formula[3]

$$\tau \simeq \frac{d^2}{D}, \tag{31}$$

where d is the transverse dimension of the sample, D is the coefficient of magnetic diffusion given by the following expression[3]

$$D = \frac{10^8}{0.4\pi k\mu}, \tag{32}$$

k is the specific electric conductivity, and μ is the magnetic permeability.

The degree of relaxation due to the diffusion of magnetic flux is determined by the following expression[3]

$$\Delta_0 = 4\pi\mu^{-1}\lambda^2 E_\infty, \tag{33}$$

where $\lambda = \partial\epsilon/\partial H$ is the tension strain produced by a change in the intensity of the field. As Becker and Döring[177] have shown, the degree of relaxation Δ_0 has its maximum value at a magnetization that is close in magnitude to the remanent magnetization; in this case

$$(\Delta_0)_{max} \cong 0.3\lambda_s \frac{E_\infty}{\sigma_i}, \tag{34}$$

where λ_s is the tension strain, which corresponds to a change in magnetization from zero to saturation, while σ_i is a measure of the remaining microstresses which may be due to previous plastic deformations or to the presence of impurity atoms.

Losses of the second kind are due to the fact that each domain reacts to the deformation produced by the oscillation of the sample, and consequently local variations of magnetic flux and local eddy currents.

The third type of losses arises when the deformations produce an irreversible displacement of domain boundaries. The relaxation times have not been determined for the second and third types of losses. For the second type of losses it has been found that[177,178]

$$Q_1^{-1} = \frac{42.6\, l^2 J_s^2 \nu^2 \sigma^2}{\pi\rho\sigma_i^2}, \tag{35}$$

where l is the periodicity of the distribution of internal stresses σ_i, J_s is the limiting magnetization, ν is the oscillation frequency, σ is the applied stress, and ρ is the density. In the case of magnetomechanical hysteresis[178]

$$Q_2^{-1} = \frac{4b\sigma E_\infty}{3\pi}, \tag{36}$$

where

$$b = \frac{d(E^{-1})}{d\sigma}.$$

The losses of elastic energy accompanying oscillations of a ferromagnetic rod may be divided into their component parts fairly simply. Indeed, the first type of losses is excluded if the material is in a demagnetized state. In order to separate losses of the second and the third type one can make use of their different dependence on the oscillation frequency.[179]

4. Internal Friction Due to Heat Conductivity

Thermal conductivity gives rise the best known relaxation mechanism in metals and alloys, first investigated in detail by Zener.[3] According to Zener, the degree of relaxation for a homogeneous isotropic solid is given by

$$\Delta_0 = E_s \frac{T\alpha^2}{\rho c_p}, \tag{37}$$

where E_s is the "adiabatic" modulus of elasticity, T is the absolute temperature, α is the coefficient of linear thermal expansion, c_p is the specific heat at constant pressure, and ρ is the density.

The relaxation time for attaining temperature equilibrium is given by the formula

$$\tau \cong \frac{d^2}{\chi}, \tag{38}$$

where d is the distance (usually the size of the sample), comparable in magnitude with the distance that must be traversed by the heat flux in order to establish temperature equilibrium, and χ is the coefficient of temperature conductivity given by the formula

$$\chi = \frac{\kappa}{\rho c_p}, \tag{39}$$

where κ is the coefficient of heat conductivity, c_p and ρ have the same meaning as before.

Zener has also investigated[3] the case when the sample is not isotropic microscopically, i.e., when it consists of separate randomly-oriented individual crystallites whose average linear dimension is d. In this case a macroscopically homogeneous stress applied to the sample will result in a microscopic inhomogeneity of the stress from one crystal to the next. This will lead to a change of temperature in the crystals and will consequently (owing to the temperature gradient in the grains) produce thermal fluxes. The relaxation time will also be given in this case by expression (38), if we take d to mean the average linear dimension of the grain. The degree of relaxation may be estimated[174] by means of formula (37), multiplied by the correction factor

$$R = \frac{\overline{(E^{-2})} - \overline{(E^{-1})}^2}{\overline{(E^{-2})}}, \tag{40}$$

which determines the relative value of the mean

square deviation of the modulus of elasticity of neighboring crystallites. The factor R shows that the degree of relaxation (and the maximum of internal friction) will increase as the elastic anisotropy of individual crystallites increases.

In going over to anisotropic solids Zener, as was first correctly noted by Isakovich,[242] makes a large simplification by taking into account only the equalization of temperatures within the boundaries of each individual crystallite and not taking into account heat exchange between them. Such a simplification leads to a change in the nature of the dependence of internal friction on frequency.

From problems of damping of oscillations of finite bodies investigated by Zener we can go over to an investigation of the absorption of sound in an unbounded medium. The physical problems associated with this, and the asymptotic estimates of the damping coefficient for sound in anisotropic unbounded media have been investigated by Landau and Lifshitz,[173] and later by Lessen[246] in a form which is more compact and tractable than Zener's treatment.

In conclusion we note that internal friction due to thermal conductivity is as a rule of small magnitude, since the degree of relaxation Δ_0, for example, in the case of α-brass[174] amounts to only 0.0036, while relaxation due to atomic diffusion in the case of the same alloy is ~ 0.5.

5. Internal Friction Due to the Interaction of the Acoustic Field with Phonons and Conduction Electrons

In the absence of an acoustic field the numbers of phonons and of electrons in any arbitrary state are determined respectively by the Planck and the Fermi distribution functions. When an acoustic field is imposed, for example, by exciting free vibrations of the sample, the distribution functions are altered because of the interaction of the sound waves with the phonons and the electrons. As a result of such interactions, the energy of acoustic vibrations is gradually distributed among all the phonons and all the electrons. The amount of the energy of sound vibrations that is dissipated depends above all on the ratio of the wavelength of sound to the mean free path of the phonons or for electrons.[188]

The scattering of sound by phonons when the wavelength of sound is considerably greater than the phonon mean free path was first investigated in references 188 to 191. For experiments carried out at temperatures above the characteristic temperature, the internal friction turned out to be proportional to the frequency of sound vibrations ν and to the temperature T. The second part of reference 190 treats the scattering of sound due to the interaction not only with phonons but also with conduction electrons. It was under the same conditions found that $Q^{-1} \sim \nu T$. If the temperature at which the experiment is carried out is $T < \theta$, then $Q^{-1} \sim \nu/T^5$.

Comparatively recently the question of the dissipation of energy of sound vibrations as a result of the interaction with conduction electrons, which are assumed to be free, has been again considered in reference 253. For the case of longitudinal waves it was found that

$$Q_1^{-1} = \frac{2Nm}{\rho\omega\tau}\left(\frac{1}{3}\frac{k^2l^2\tan^{-1}kl}{kl-\tan^{-1}kl} - 1\right). \tag{41}$$

Here N is the number of electrons per unit volume, m is the electron mass, ρ is the density of the metal, ω is the circular frequency of the sound, k is the wave number, while τ and l are the mean collision time and the mean free path for the electron, respectively. In the case of transverse waves the magnitude of internal friction is given by

$$Q_2^{-1} = \frac{2Nm}{\rho\omega\tau}\frac{1-g}{g}, \tag{42}$$

where

$$g = 3\left(2k^2l^2\,\frac{k^2l^2+1}{kl\tan^{-1}kl-1}\right)^{-1}. \tag{43}$$

In the case of low frequencies, when $\lambda \gg l$ and $kl \ll 1$, we obtain from (41) and (42):

$$Q_1^{-1} \cong \frac{8Nmv_0^2\tau}{15\rho v_l^2}, \tag{44}$$

$$Q_2^{-1} \cong \frac{2Nmv_0^2\omega\tau}{5\rho v_s^2}, \tag{45}$$

where v_0 is the electron velocity, while v_l and v_s are the velocities of the longitudinal and the transverse waves respectively.

In the case of high frequencies, when $\lambda \approx l$ and $kl \approx 1$,

$$Q_1^{-1} \cong \frac{2\pi Nmv_0}{6\rho v_l} \tag{46}$$

and

$$Q_2^{-1} \cong \frac{16Nmv_0}{3\rho v_s\omega}. \tag{47}$$

The last condition is fulfilled in the region of ultrahigh frequencies ($\nu \sim 10^6$ cps) at very low temperatures, when the electron mean free path becomes comparatively large. Experimental work[250,251] on the dissipation of ultrasound in pure

metals at very low temperatures (Figs. 2 and 3) is in satisfactory agreement with the theory. However, Pippard's theory[253] does not explain the sharp falling off in the absorption sound when the metal goes over into the superconducting state (Fig. 2, curve V). An attempt to treat this interesting problem was made by Mason.[252]

At very high frequencies ($\nu \sim 10^9$ cps) the interaction of phonons with conduction electrons plays an essential role. This interaction has been taken into account by Paranjape,[192] who found for the case of longitudinal waves:

$$Q^{-1} = 4\left(\frac{\pi}{6}\right)^{\frac{1}{3}} \frac{v_l n^2 n_0^{\frac{2}{3}} e^2}{k} \frac{\partial \sigma}{\partial T}; \qquad (48)$$

here ν_l is the velocity of sound, n is the number of electrons in an atom, n_0 is the number of atoms per unit volume, e is the electronic charge, k is the wave number, and σ is the specific electric resistance. In this case, too, the dissipation of energy turns out not to depend on the frequency. It is not possible to check this theory since experiments at such frequencies have not yet been performed.

6. Internal Friction Due to the Migration of Atoms in a Stress Field

An essential role may also be played by irreversible processes other than relaxation.. Some of these have been considered in the preceding section. In Secs. 6 and 7 we shall extend our investigation of the dissipation of energy due to various irreversible processes.

As is well known, the crystal lattice of a real crystal may possess various types of imperfections (defects). These imperfections appear either during crystallization from a melt, or as the result of a preceding plastic deformation, and also of thermal motion of atoms.[194] The presence of defects produces continuous displacements of atoms in the crystalline lattice. Each transition of an atom is associated with overcoming a certain potential barrier u_i, whose magnitude depends on the type of the transition and on the state of the sample as a whole. The probabilities of individual transitions are proportional[195] to $e^{-u_i/kT}$

On considering various transitions of atoms within the lattice of the sample we find[137] that the total number of different transitions per unit time per unit volume is

$$w' = w_0 e^{-u/kT}. \qquad (49)$$

In an unstressed state such transitions occur at random and in such a way that the mean thermal energy of the crystal at constant temperature remains constant. This means that there is no spontaneous dissipation of energy by a crystal which is in thermal equilibrium with the surroundings.

In a deformed sample each volume element of the sample will experience a shearing stress $\theta(t)$ (for example, in torsional oscillations). In this case the number of irreversible transitions leading to the dissipation of elastic energy will be determined by the difference between the total number of transitions in the presence of the stress field and the number of transitions which are caused by thermal fluctuations in the absence of the field.

Therefore, the number of irreversible transitions per unit time due to the presence of the stress field is equal to

$$\Delta w = w''\left(\frac{u}{kT} - \frac{\beta\bar{\theta}}{kT}\right) - w'\left(\frac{u}{kT}\right), \qquad (50)$$

where $\beta\bar{\theta}$ is the average change in the potential barrier u caused by the shearing stress $\bar{\theta}$. Since these stresses are small we have $\beta\bar{\theta} \ll kT$, and the parameter $\beta\bar{\theta}/kT$ will be small compared with unity. On expanding ω'' into a Taylor's series in powers of the small parameter $\beta\bar{\theta}/kT$, and on restricting ourselves to first-order terms, we obtain

$$\Delta w \cong \frac{\beta\bar{\theta}}{kT} w_0 e^{-u/kT}. \qquad (51)$$

Introducing the mean "activation" energy H per mole, we obtain

$$\Delta w \cong \frac{\beta\bar{\theta}}{kT} w_0 e^{-H/RT}. \qquad (52)$$

The above expression for $\Delta\omega$ may be easily related to the magnitude of the internal friction Q^{-1} for this process. Indeed, the relative energy $\Delta U/U$ dissipated per cycle of oscillations is equal to $(a_n^2 - a_{n+1}^2)/a_n^2$ where a_n and a_{n+1} are two successive amplitudes separated by a time interval equal to one period. The energy dissipated per unit time will be larger by a factor ν. Assuming that this energy is proportional to the number of irreversible transitions Δw, we obtain

$$\frac{a_n^2 - a_{n+1}^2}{a_n^2} = \alpha \Delta w.$$

On taking account of the fact that $Q^{-1} = \dfrac{\ln \dfrac{a_n}{a_{n+1}}}{\pi}$, we obtain

$$\Delta w = \frac{\nu}{\alpha}(1 - e^{-2\pi Q^{-1}}).$$

Since $Q^{-1} < 0.1$, we can expand the exponential

within the brackets of the preceding expression into a series in powers of $2\pi Q^{-1}$ and restrict ourselves to the first two terms of the expansion. Taking (52) into account, we then obtain[93,137]

$$Q^{-1} = \frac{\alpha\beta\bar{\theta}w_0}{\omega kT} e^{-H/RT}. \quad (53)$$

From this expression for Q^{-1} it can be seen that internal friction is inversely proportional to the frequency ω and directly proportional to the average value of the external shearing stress $\bar{\theta}$. Moreover, internal friction will decrease as the "activation energy" H, characterizing the possibility of various types of atomic transitions within the lattice, increases. The internal friction determined by (53) increases with increasing temperature, reaching a maximum at a temperature

$$T_m = \frac{H}{R}, \quad (54)$$

whose magnitude depends on the value of H.

7. Internal Friction Due to Static Hysteresis

In contrast to dynamic hysteresis[53] which occurs in an oscillating sample when the deformation is not a single valued function of the stress, hysteresis can also occur in several solids when they are statically stressed. Such a possibility arises only in that case when there remains after the removal of the stress a constant remanent deformation that can be reduced to zero only by applying a stress in the opposite direction. When the stress is varied between the limits $\pm\sigma$ the result is a static hysteresis loop whose area characterizes the energy losses (similar to the magnetomechanical hysteresis discussed earlier).

Static hysteresis is possible principally[196] in the case of stresses that lead to plastic deformation. Therefore the latter are of importance in the study of fatigue and are not very important in the study of internal friction at small amplitudes. However, it has been shown[84,197] comparatively recently that the damping of oscillations in single crystals at deformations on the order of 10^{-7}, a damping that can be ascribed to the motion of dislocations, is also accompanied by static hysteresis. Thus, static hysteresis is possible also in the case of very small amplitudes as a result of a particular kind of atomic rearrangement (or of magnetic rearrangement in the case of magnetoelastic hysteresis). The speed of such rearrangements is of the order of the speed of sound, and they can therefore be considered to be practically "instantaneous" in the case of low-frequency measurements of internal friction. In contrast to the relaxation type of internal friction, static hysteresis shows a noticeable amplitude dependence even at small amplitudes.

In the case of static hysteresis the internal friction is given by the following formula[4,53]

$$Q^{-1} = \frac{E_\infty}{\pi}\left[(A_0 - a_0) + \frac{2}{3}(A_1 - a_1)\sigma_1 + \frac{1}{2}(A_2 - a_2)\sigma_2^2 + \dots\right], \quad (55)$$

where E_∞ is the unrelaxed modulus of elasticity, A_0, A_1, A_2 are the coefficients in the expansion of $d\epsilon''/ds = F(\sigma)$ (as the load is applied) into a power series, a_0, a_1, a_2 are the coefficients in the expansion of $d\epsilon''/d\sigma = f(\sigma)$ (as the load is removed) into a power series; ϵ'' is the inelastic part of the total deformation ϵ; $\sigma = \sigma_i \cos \omega t$ is the periodically varying stress which appears in the sample when the load is applied, and σ_i is the amplitude of this stress.

In the case of a loop which has the form of a parallelogram in which F and f are constant, the magnitude of the internal friction Q^{-1} will not depend on the stress amplitude σ_i. In general, if there are no grounds for assuming such a shape for the loop, the existence of static hysteresis is closely related to the amplitude-dependent dissipation of energy even at very small amplitudes. This characteristic amplitude dependence of Q^{-1}, together with its independence of the frequency in the case of static hysteresis, makes it possible to distinguish relaxation phenomena from phenomena belonging to this class.

8. Internal Friction and Dislocations

The internal friction of a well annealed material is usually of the order 10^{-5}, although its magnitude depends on the purity of the metal and on the amplitude of the oscillations. However, this minimum of internal friction also depends strongly on the experimental conditions, particularly in the case of single crystals. In references 42, 199, and 200 it was reported that internal friction increases strongly as a result of weak shocks, and even as a result of different manipulations of the crystal. Moreover, single crystals prepared from the same material and by the same method differ sometimes by more than a factor ten with respect to their internal friction.[201,202] Even different parts of the same crystal have unequal internal friction.[197] An even greater scattering of the values of internal friction is observed in the case of materials that had been previously deformed.[24,32,38,42,85,90,94,96,98,100,101,108,114,118,128,199,200,202-225]

The amplitude dependence of internal friction observed in the case of annealed materials and of

"cold worked" materials is difficult to explain within the framework of the relaxation theory of internal friction. In connection with this, attempts have been made recently to utilize the theory of dislocations to explain the phenomena mentioned above.[24,80,84,147,198,214,216,222,227-232] The dislocation theory of internal friction is based on Koehler's model.[229] According to this model the dissipation of energy is due to the oscillations under the influence of a periodic external force of the dislocation loop associated with impurities.

It is assumed that the pure single crystal contains a network of dislocations prior to the deformation. When an external stress is applied a deformation due to dislocations takes place in addition to the elastic deformation, and is called a deformation dislocation. The qualitative behavior of a dislocation under the influence of an increasing external force is shown[232] in Fig. 25. In the case of zero stress the length L_N of the dislocation is linked by impurity particles (Fig. 25, A). For very small stresses (B) the loops L_c sag and continue to sag until the disruptive stress has been reached (C-D). As the stress is increased further the total length of the dislocation increases (D-F), until finally a closed dislocation loop (G) is formed. Because of this process, the dislocation deformation must be irreversible and can therefore be classed as a plastic deformation.

A consistent examination of this mechanism leads to the conclusion that losses may be of two types: those that depend on the frequency of oscillation of the sample, and those that do not. In case of losses of the first type the internal friction can be represented by the formula[232]

$$Q_1^{-1} = \frac{\Delta_0 \Lambda L^2}{2\pi D}\left[\frac{\Omega}{(1-\Omega^2)^2 + \frac{\Omega^2}{D^2}}\right], \qquad (56)$$

where Λ is the total length of the moving dislocation lines, L is the length of the dislocation loop,

$$\Delta_0 = \frac{8\mu a^3}{\pi^2 c},$$

μ is the shear modulus, a is Burgers' vector,

$$D = \frac{\omega_0 B}{A}, \quad \Omega = \frac{\omega}{\omega_0},$$

ω_0 is the resonance frequency, and ω is the frequency of oscillation of the sample. The quantities A, B, and C are the coefficients in the equation of motion of bound dislocations:

$$A\frac{\partial^2 \xi}{\partial t^2} + B\frac{\partial \xi}{\partial t} - C\frac{\partial^2 \xi}{\partial y^2} = a\sigma,$$

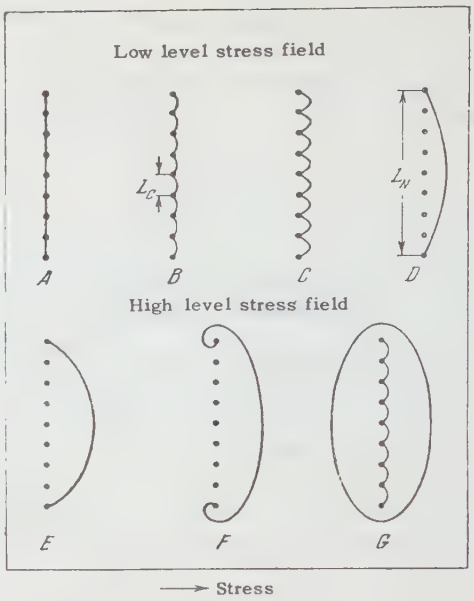


FIG. 25. Schematic diagram of the formation of dislocation loops in a stress field.

where $\xi(x, y, t)$ is the displacement of an element of the dislocation loop from its equilibrium position, y is the coordinate of this element; σ is the stress.

In the case of losses of the second type the internal friction is given by

$$Q_2^{-1} = \frac{\Delta_0 \Lambda L_N^2}{2\pi^2}\left(\frac{L_N}{L_c}\right)\left[\frac{\Gamma}{\sigma_0} - 1\right] e^{-\frac{\Gamma}{\sigma_0}}, \qquad (57)$$

where $\Gamma = L_b\sigma/L_c$, L_b is the length of the dislocation, $\sigma = \sigma_0 \cos(\omega t - kx)$ is the stress, and the remaining quantities have their previous meaning.

The theory presented briefly above does not apply to high-temperature regions. The first attempt to solve the problem of high-temperature internal friction from the point of view of dislocation theory was made by Weertman.[243] His calculation is based on the results of the theory of microcreep[244,245] and of the theory of quenching.[244] Assuming that a low amplitude stress causes the dislocations to move in the glide plane Weertman obtains for the internal friction an approximate equation of the following form:

$$Q_1^{-1} = Q_0^{-1} e^{-H/kT}, \qquad (58)$$

where the factor in front of the exponential is given by

$$Q_0^{-1} = \frac{3\alpha\mu b^5 N \nu_0}{2\pi 4^{4/3} \nu kT c^{13/9} \beta^{2/3}\left(\ln\frac{1}{c}\right)^{2/3}}, \qquad (59)$$

and the heat of activation is given by

$$H = \frac{3}{4^{2/3}\pi} b^3 c^{\frac{1}{9}} \mu \beta^{\frac{2}{3}}\left(\ln\frac{1}{c}\right)^{\frac{2}{3}}. \qquad (60)$$

The last two expressions contain the following quantities: α is a constant equal to $1/25$, μ is the shear modulus, b is the distance between nearest neighboring atoms, N is the density of dislocations, ν_0 is the frequency of oscillation of a dislocation line in its potential well, ν is the frequency of the oscillation of the sample, c is the impurity concentration, and $\beta = (1/a)(da/dc)$ where a is the lattice constant of the alloy. A displacement of a dislocation in the stress field created by the external periodic force may take place not only in the glide plane.

By examining other possibilities for the displacement of dislocations in a stress field, Weertman obtains another expression for the magnitude of internal friction at high temperatures:

$$Q_2^{-1} = \frac{\alpha \mu b^5 \nu_0^* N}{2\pi \nu kT} e^{\frac{H^* - ST}{kT}}, \qquad (61)$$

where ν_0^* is the frequency of oscillation of a vacancy, H^* is the activation energy for self diffusion, and S is the activation entropy for self diffusion. The other symbols remain the same.

The second mechanism proposed by Weertman can also be applied to low-frequency oscillations, but again only for small amplitudes, $\sim 10^{-7}$. It is well known that all the low frequency methods ($\nu \sim 1$ cps) utilize oscillation amplitudes $\sim 10^{-5}$ and greater, which excludes the possibility of applying Weertman's theory to the numerous experimental results obtained by these methods.

A comparison of the theory with the experimental results obtained by high-frequency methods leads, for example in the case of aluminum, to the following results.

The theory yields $H = 11{,}000$ cal/g at., $\Delta_0 \cong 270$; while calculations based on two points of the experimental curve give $H = 12{,}900$ cal/g at., $\Delta_0 = 100$. The agreement between theory and experiment in this case is thus quite satisfactory. The degree of success in explaining internal friction by means of the theory of dislocations has so far not been very great, because internal friction of deformed samples is a very complicated phenomenon that depends on many factors: on the degree and nature of preceding deformation, on the time of aging, on the presence of impurities in the sample, on the frequency of oscillations, on the temperature at which measurements are carried out, etc. To attain a greater degree of success it is necessary to carry out systematic investigations, taking into account many parameters that affect internal friction, particularly at low frequencies, when static hysteresis is observed.

IV. APPLICATION OF THE THEORY TO THE EXPLANATION OF THE TEMPERATURE DEPENDENCE OF INTERNAL FRICTION OF METALS AND ALLOYS

From a survey of experimental data it is easily seen that the problem of temperature dependence of internal friction has been studied mostly at temperatures above 0° C and largely by low-frequency methods.

We shall therefore utilize the theoretical picture of internal friction developed above to explain the temperature dependence $Q^{-1}(T)$ of pure metals and alloys in the case of low-frequency oscillations ($\nu \sim 1$ cps). Of the pure metals we shall select for this purpose aluminum, the metal most investigated in all respects, and of the alloys we shall take, for example, Ni-Fe. Earlier we have said that if the total internal friction does not depend on the time and is small ($Q^{-1} < 0.1$), then it can be expressed as the sum of contributions made by the different dissipative mechanisms acting independently of one another. In computing the total value of internal friction it is necessary to take into account the constant contribution due to the various losses in the apparatus. This part of the losses depends on the construction of the apparatus and on the method of measurement. However, it does not exceed the value $\sim 1 \times 10^{-4}$. Thus

$$Q^{-1} = Q_a^{-1} + \sum_{i=1}^{n} Q_i^{-1}, \qquad (62)$$

where Q_i^{-1} are the contributions made by the various dissipation mechanisms. Let us make an estimate of the magnitude of the various contributions for the case of low frequency torsional oscillations of an aluminum sample. In references 93 and 137 such an estimate has been made for various pure metals. In particular, in the case of a single crystal of pure aluminum (99.98% Al), an estimate of the contributions of the various dissipative mechanisms discussed above yields temperature dependence of internal friction in the form

$$Q_{Al}^{-1} = 1.1 \cdot 10^{-5} + 4.5 \cdot 10^{-6} t, \qquad (63)$$

where t is the temperature in °C at which the measurements were made.

In Fig. 26 this quantity is represented by the straight line III which coincides with the experimetal curve I for the internal friction of a single crystal aluminum sample only up to 240° C. The reason for the discrepancy, as we shall see later, is the imperfection of real single crystals. From the same diagram it can also be seen that the internal friction of polycrystalline aluminum (II) differs from the internal friction of single-crystal

aluminum (I). The appearance of a maximum on curve II may be explained on the basis of the concept of a "two-component system" discussed earlier. In the present case one of the components will be the grain (assumed to be elastic), while the other component will be the grain boundary (assumed to be viscous). When such a system undergoes periodic deformation, displacements along the grain boundaries take place and lead to dissipation of energy. This mechanism in the case of aluminum was discussed in detail for the first time in reference 15 and again in references 93, 137, and 224. Using the known values of Δ_0 and τ for polycrystalline aluminum, an evaluation was made, with the aid of formula (15), of the contribution to the internal friction due to the shift in the grain boundaries, on the assumption that this mechanism can be described by one relaxation time τ and one heat of activation H. The calculations gave the curve IV shown in Fig. 26, which differs strongly from the experimental curve II. This discrepancy can be partially eliminated by assuming that a displacement along the grain boundaries corresponds to a spectrum of relaxation times and to a set of heats of activation with the maximum density in the neighborhood of a certain mean value. By making a suitable choice of τ_i in the neighborhood of τ_{av}, of $(\mu_\infty/\mu_0)_{av}$, and of H_{av} it is possible to obtain the curve V. This curve taken together with the straight line III still does not give a complete description of the experimental curve II, particularly at high temperatures. But at high temperatures the dissipation mechanism associated with migration of atoms in the stress field becomes important. Its contribution is described by formula (57). In references 93 and 137 the following value for this contribution was obtained

$$Q_1^{-1} = \frac{1.13 \cdot 10^4}{T} e^{-\frac{8,900}{RT}}, \qquad (64)$$

which is shown in Fig. 26 by the curve VI. Now the sum of all the contributions given by the curve VII agrees well with the experimental curve II.

Returning to the internal friction of single-crystal aluminum, we can explain the discrepancy between the experimental curve of internal friction (I) with the straight line for the internal friction (III) by the fact that the dissipation of energy due to various transitions of atoms in the lattice in the stress field takes place also in the single-crystal state of the material. The internal friction in the case of this mechanism is given by[93,137]

$$Q_2^{-1} = \frac{7.12 \cdot 10^4}{T} e^{-\frac{13,800}{RT}}. \qquad (65)$$

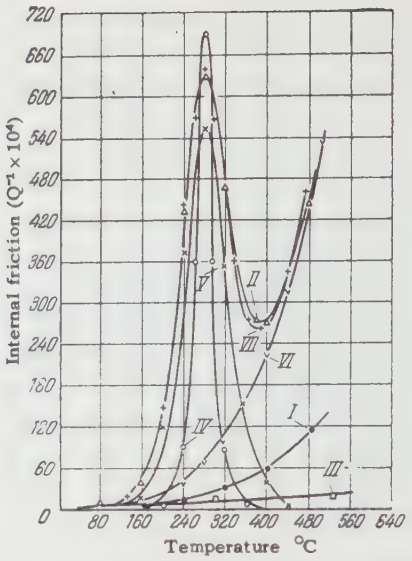


FIG. 26. Temperature dependence of internal friction in aluminum: I and II are experimental curves; III – VII are theoretical curves.

By comparing the quantity Q_1^{-1} for polycrystalline aluminum [formula (64)] with the value of internal friction Q_2^{-1} for single-crystal aluminum [formula (65)], we see that the parameter H is considerably larger for the single-crystal state of aluminum than for the polycrystalline state. Such a result does not contradict the modern concept of the single-crystal state of the metal as being a more perfect state (in the sense of the degree of order in the position of the particles), than the polycrystalline state. A similar satisfactory result was obtained[93,137] for the internal friction $Q^{-1}(T)$ of copper, nickel, iron, and cobalt (at high temperatures, since at intermediate temperatures $Q_{c_0}^{-1}$ is not constant because of the allotropic transitions $C_{0\beta} \rightleftharpoons C_{0\alpha}$).

In our earlier considerations of the temperature dependence of internal friction of alloys we arrived at the conclusion that, firstly, the internal friction of alloys is considerably smaller than the internal friction of pure metals of which the alloy is composed almost over the whole range of temperatures; secondly, the maximum in the internal friction due to viscous flow along grain boundaries often degenerates into an inflection point as the concentration of dissolved atoms is increased; thirdly, in the low temperature region there appear maxima which are absent in the case of pure metals. Since the nature of the internal friction of alloys changes gradually as the concentration of the dissolved atoms increases,[103,104,137] we can say that the internal friction of alloys can also be represented (with a few exceptions[233]) as a sum of contributions Q_i^{-1}.

By analyzing the different contributions Q_i^{-1},

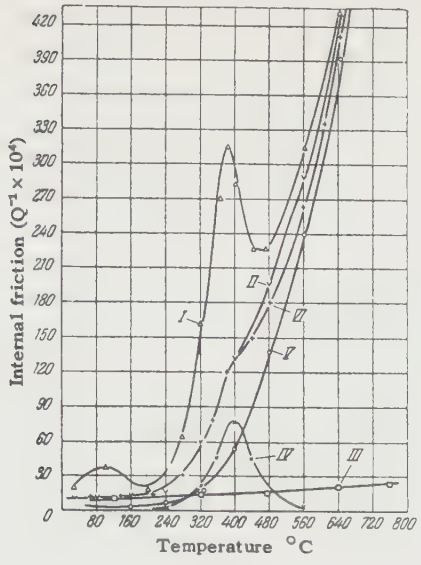


FIG. 27. Temperature dependence of internal friction in nickel (I) and in a solid solution of iron (2.6% Fe) in nickel. I and II are experimental curves; III – VI are theoretical curves.

for example, for Ni-Fe[137] in a manner similar to that employed in the case of pure metals, and by representing the results in graphical form, we obtain Fig. 27. In this diagram curve I represents the temperature dependence of internal friction of pure nickel (99.99% Ni), and curve II represents the temperature dependence of internal friction of a solid solution of iron in nickel. The straight line III represents the internal friction associated with all the processes that give a linear dependence of Q^{-1} on the temperature:

$$Q_{III}^{-1} = 8.5 \cdot 10^{-4} + 2.27 \cdot 10^{-6} t. \qquad (66)$$

Curve IV represents that part of the internal friction which is due to the viscous slipping along the grain boundaries, and which is determined by formula (16); the mean heat of activation for such a process is ~ 60,000 cal/mole. Curve V represents the contribution due to the migration of atoms in the stress field; its value is

$$Q_V^{-1} = \frac{28.2 \cdot 10^3}{T} e^{-\frac{12{,}000}{RT}}. \qquad (67)$$

Finally, curve VI gives the sum of all these contributions and, as can be seen, it quite satisfactorily reproduces the shape of the experimental curve II. For pure nickel the mean heat of activation of the process of relaxation of stress along the grain boundaries is 63,000 cal/mole, while the value of the maximum for this contribution is $\sim 165 \times 10^{-4}$. The contribution due to the migration of atoms is given by

$$Q_{Ni}^{-1} = \frac{2.76 \cdot 10^3}{T} e^{-\frac{8200}{RT}}. \qquad (68)$$

On intercomparing the results obtained above for the internal friction of pure nickel and of the solid solution of iron in nickel (similar results have also been obtained for many other metals and alloys[137]) we see that a reduction in the internal friction of alloys in comparison with pure metals is due to the decrease, in the first instance, of the magnitude of the contributions associated with viscous slipping along grain boundaries, and with the migration of atoms in the stress field (through an increase in the average "activation" energy H for the various types of transitions of the atoms in the alloy lattice).

The low temperature maxima occurring in the case of alloys and of certain not quite pure metals (iron, tantalum, and others), are associated with the presence of foreign atoms (such as hydrogen, carbon, nitrogen, boron, oxygen, etc.), and are due primarily to the preferential distribution of the atoms in the stress field.[152–166]

CONCLUSION

The material presented above enables us to say that the physical picture of the dissipation of elastic energy in oscillating samples is sufficiently clear. The existing theories describe (qualitatively, and in certain cases, also quantitatively) in a satisfactory manner the temperature dependence of internal friction of well-annealed metals and alloys which do not undergo phase transitions on heating. First attempts have been made to describe the behavior of internal friction of plastically deformed metals and alloys. However, the theory of internal friction, in view of the complexity of the phenomenon of energy dissipation itself, and of the small number of reliable systematic investigations (particularly of plastically deformed metals and alloys), which take into account the fact that internal friction always depends on many factors, is still far from complete. There are very few investigations (either theoretical or experimental) of internal friction at high temperatures. And yet investigations of internal friction at high temperatures provide us with still another possibility of gaining an understanding of the problem of the behavior of alloys at high temperatures which should not be neglected.

[1]B. Hopkinson, G. T. Williams, Proc. Roy. Soc., Ser. A, 87, 502 (1912).

[2]O. Föppl, J. Iron a. Steel Inst. 134, 393 (1936).

[3]C. Zener, Elasticity and Anelasticity of Metals, 1948.

[4]A. S. Nowick, Progress in Metal Physics 4, 1 (1953).

[5] C. Zener, J. Appl. Phys. 18, 1022 (1947).
[6] T. S. Kê, M. Ross, Rev. Scient. Instr. 20, 795 (1949).
[7] W. Weber, Poggendorf's Ann. 34, 247 (1937).
[8] K. Jokibe, S. Sukai, Phil. Mag. 42, 397 (1921).
[9] E. A. Kopilovich, Sow. Phys. 3, 561 (1933).
[10] A. Gemant, W. Jackson, Phil. Mag. 23, 960 (1937).
[11] F. Förster, H. Breitfeld, Z. Metallkunde 30, 343 (1938).
[12] E. B. Lunts, Прикладная математика и механика (Appl. Math. and Mech.) 1, 331 (1938)
[13] J. Norton, Rev. Sci. Instr. 10, 77 (1939).
[14] P. Chevenard, Rev. Met. 39, 65 (1942).
[15] T. S. Kê, Phys. Rev. 71, 533 (1947).
[16] Ch. Boulanger, Rev. Met. 46, 321 (1949).
[17] R. Kamel, Phys. Rev. 75, 1606 (1949).
[18] C. A. Wert, J. Appl. Phys. 20, 29 (1949).
[19] M. E. Fine, J. Metals 2, 1322 (1950).
[20] F. K. Gorskii, J. Tech. Phys. (U.S.S.R.) 20, 1111 (1950).
[21] L. Glikman and K. Shishokina, Завод. лаб. (Plant Laboratory) 11, 1365 (1951).
[22] V. I. Prosvirin and N. N. Morgunova, Труды ЦНИИТМАШ (Trans. of Central Scientific Research Institute for Technology and Machine Building) No. 40 (1951).
[23] J. W. Jensen, Rev. Scient. Instr. 23, 397 (1952).
[24] R. E. Maringer, J. Appl. Phys. 24, 1525 (1953).
[25] C. C. Lo and M. Gensamer, Acta Metallurgica 1, 483 (1953).
[26] V. S. Postnikov, Труды Кемеровского горного института (Trans. Kemerovo Mining Inst.) 1, 211 (1954).
[27] K. Misek, Чех. физ. журнал (Czech. Phys. J.) 4, 67 (1954).
[28] S. O. Tsobkallo and V. A. Chelnokov, J. Tech. Phys. (U.S.S.R.) 24, 499 (1954).
[29] C. Ang and C. Wert, J. Appl. Phys. 25, 1061 (1954).
[30] B. S. Berry, Rev. Scient. Instr. 26, 884 (1955).
[31] V. S. Postnikov and Yu. V. Piguzov, ИТЭИН АН СССР (Inst. Tech. Inform., Acad. Sci. U.S.S.R.) "Tools and Stands" PS - 55 - 448.
[32] G. B. Brook and A. H. Sully, Acta Metall. 3, 460 (1955).
[33] S. Weinig, Rev. Scient. Instr. 26, 91 (1955).
[34] S. L. Quimby, Phys. Rev. 25, 558 (1925).
[35] J. Zacharias, Phys. Rev. 44, 116 (1933).
[36] L. Poltavskii, Sow. Phys. 3, 555 (1953).
[37] W. T. Cooke, Phys. Rev. 50, 1158 (1936).
[38] R. L. Wegel and H. Walther, Physics 6, 141 (1936).
[39] D. Bencroft and R. B. Jacobs, Rev. Scient. Instr. 9, 279 (1938).
[40] R. H. Randall, F. C. Rose, and C. Zener, Phys. Rev. 56, 343 (1939).
[41] A. Gemant, J. Appl. Phys. 11, 647 (1940).
[42] A. W. Lawson, Phys. Rev. 60, 330 (1941).
[43] T. S. Kê, J. Appl. Phys. 20, 1226 (1949).
[44] P. G. Bordoni, Nuovo cimento 7, 162 (1950).
[45] A. V. Siefert and F. T. Worrell, J. Appl. Phys. 22, 1257 (1951).
[46] J. K. Páttison, Rev. Scient. Instr. 25, 67 (1954).
[47] K. Tanaka, H. Abe, and K. Hirano, J. Phys. Soc. Japan 10, 454 (1955).
[48] R. Cabarat, Akust. Beihefte 1, 200 (1956).
[49] C. Peterson, Arch. techn. Messen 164, 74 (1949).
[50] E. Diepschlag and H. Müller, Arch. Metallkunde 3, 400 (1949).
[51] E. S. Sorokin, Central Scientific Research Institute for Commercial Structures No. 15 (1954).
[52] G. Kol'skii, Волны напряжения в твердых телах (Stress Waves in Solids) GIIL, 1955, Ch. V, VI.
[53] G. Löschner, Technik 11, 289 (1956).
[54] V. D. Kuznetsov, Физика твердого тела (The Physics of Solids) vol. 4, Ch. 30, 1947.
[55] C. Barus, Phil. Mag. 29, 337 (1890).
[56] R. Voigt, Ann. d. Phys. 47, 671 (1892).
[57] W. Heydweiber, Wied. Ann. 63, 56 (1897).
[58] Reiger, Phys. Z. 2, 213 (1901).
[59] B. P. Veinberg, J. Russ. Phys. Soc. 36, 47 (1904).
[60] Trouton and Rankine, Phil. Mag. 8, 538 (1904).
[61] B. P. Veinberg, Записки Новосибирского университета (Notes, Novosibirsk University) 105, 157 (1905).
[62] B. P. Veinberg, J. Russ. Phys. Soc. 44, 201 (1912).
[63] B. P. Veinberg, J. Russ. Phys. Soc. 45, 701 (1913).
[64] K. Honda, and S. Konno, Phil. Mag. 42, 115 (1921).
[65] B. P. Veinberg, Изв. Томского техн. ин-та (Bull. Tomsk Tech. Inst.) 43, 39 (1923).
[66] V. D. Kuznetsov, Изв. Томского техн. ин-та (Bull. Tomsk Tech. Inst.) 43, 47 (1923).
[67] B. P. Veinberg, Изв. Томского техн. ин-та (Bull. Tomsk Tech. Inst.) 45, 27 (1924).
[68] V. D. Kuznetsov and L. A. Shvirk, Сообщен. о научно-технических работах в республике (Communications on the Scientific and Technical Work in the Republic) 14, 1 (1924).
[69] V. D. Kuznetsov, Изв. Томского техн. ин-та (Bull. Tomsk Tech. Inst.) 46, 42 (1925).
[70] V. D. Kuznetsov and L. A. Shvirk, Изв. Томского государственного университета (Bull. Tomsk State Univ.) 75, 99 (1925).
[71] V. D. Kuznetsov, J. Russ. Phys. Soc. 56, 295 (1928).

[72] F. Hettwer, Wien. Ber., IIa, **134**, 51 (1925).
[73] G. Subrahmaniam and D. Gunnaija, Phil. Mag. **49**, 711 (1925).
[74] G. Subrahmaniam, Phil. Mag. **50**, 716 (1925).
[75] V. D. Kuznetsov, Fifth Congress of Russian Physicists, Moscow, 1926, p. 24.
[76] V. D. Kuznetsov, J. Russ. Phys. Soc. **60**, 339 (1928).
[77] B. Ya. Pines, J. Russ. Phys. Soc. **59**, 157 (1927).
[78] V. S. Postnikov, Usp. Fiz. Nauk **53**, 87 (1954).
[79] T. Kamoto, Nature **175**, 948 (1955).
[80] P. G. Bordoni, Ric. Scient. **19**, 851 (1949).
[81] P. G. Bordoni, J. Acoust. Soc. **26**, 495 (1954).
[82] J. Weertman, E. J. Salkovitz, Acta Metallurgica **3**, 1 (1955).
[83] W. B. Nowack, Phys. Rev. **85**, 710 (1952).
[84] A. S. Nowick, Phys. Rev. **80**, 249 (1950).
[85] D. H. Niblett and J. Wilus, Phil. Mag. **1**, 415 (1956).
[86] L. Rotherham, A. Smith, and B. and G. Grunough, J. Inst. Met. **79**, 439 (1951).
[87] P. G. Bordoni, J. phys. et radium **16**, 285 (1955).
[88] J. Barducci, Ric. Scient. **24**, 2025 (1954).
[89] H. Barnes, C. Zener, Phys. Rev. **58**, 87 (1940).
[90] E. G. Stanford, Research. **6**, 21 (1953).
[91] E. G. Stanford, Rev. Met. **51**, 674 (1954).
[92] H. Birnbaum and M. Levy, Acta Metall. **4**, 84 (1956).
[93] V. S. Postnikov, J. Tech. Phys. (U.S.S.R.) **24**, 1599 (1954).
[94] F. Förster and W. Köster, Z. Metallk. **29**, 116 (1937).
[95] K. J. March, Acta Metall. **2**, 530 (1954).
[96] W. Köster, L. Baugert, and W. Lang, Z. Metallkunde **46**, 84 (1955).
[97] W. Köster, Z. Metallkunde **39**, 9 (1948).
[98] A. Ya. Samoĭlova and V. S. Postnikov, Труды Кемеровского пединститута (Trans. Kemerovo Pedagogical Inst.) **3** (1957) (in press).
[99] R. March and D. Hall, Trans. AIME **197**, 937 (1953).
[100] W. Köster, L. Baugert, and J. Hafner, Z. Metallkunde **47**, 224 (1956).
[101] A. Ya. Samoĭlova and V. S. Postnikov, Dokl. Akad. Nauk S.S.S.R. **114**, 1228 (1957).
[102] J. N. Pratt, W. J. Bratina, and B. Chalmers, Acta Metallurgica **2**, 203 (1954).
[103] V. S. Postnikov, Труды Кемеровского пединститута (Trans. Kemerovo Pedagogical Inst.) **I**, 191 (1956).
[104] V. S. Postnikov, Физика металлов и металловедение (Phys. of Metals and Metal Res.) **4**, 344 (1957).
[105] T. S. Kê, C. T. Tsien, and K. Misek, Scientia Sinica **4**, 519 (1955).
[106] T. S. Kê, C. T. Tsien, and K. Misek, Cs. Cas. Fys. **6**, 37 (1956).
[107] W. Köster, Z. Metallkunde **39**, 1 (1948).
[108] T. S. Kê, Met. Techn. **15**, 2370 (1948).
[109] T. S. Kê, Scientia Sinica **4**, 55 (1955).
[110] A. Ia. Samoilova and V. S. Postnikov, Труды Кемеровского пединститута (Trans. Kemerovo Pedagogical Inst.) **3** (1957) (in press).
[111] T. S. Kê, Phys. Rev. **74**, 9 (1948).
[112] T. S. Kê, Phys. Rev. **74**, 914 (1948).
[113] W. J. Bratina and W. C. Winegard, J. Metals **8**, 186 (1956).
[114] W. Körter and K. Posenthal, Z. Metallkunde **30**, 345 (1938).
[115] C. Zener, D. Winule, and H. Nielsen, Trans. AIME **147**, 98 (1942).
[116] C. Zener, Trans. AIME **152**, 122 (1943).
[117] R. Artman, J. Appl. Phys. **23**, 475 (1952).
[118] W. Köster, Arch. Eisenhüttenw. **14**, 271 (1940).
[119] R. Hasiguti and G. Kamoshita, J. Phys. Soc. Japan **9**, 646 (1954).
[120] E. Schell and G. Reinacher, Z. Metallkunde **36**, 63 (1944).
[121] W. Köster, Z. Metallkunde **35**, 246 (1943).
[122] L. Rotherham and S. Pearson, J. Metals **8**, 881 (1956).
[123] A. S. Nowicu, Phys. Rev. **88**, 925 (1956).
[124] A. V. Grin' and V. A. Pavlov, Физика металлов и металловедение (Phys. of Metals and Metal Res.) **4**, 103 (1957).
[125] A. V. Grin', Физика металлов и металловедение (Phys. of Metals and Metal Res.) **4**, 561 (1957).
[126] A. I. Iamshchikova, Thesis, Moscow Steel Inst., 1954.
[127] A. C. Damasu and A. S. Nowicu, J. Appl. Phys. **26**, 1165 (1955).
[128] T. S. Kê, Phys. Rev. **78**, 420 (1950).
[129] T. S. Kê, J. Appl. Phys. **21**, 414 (1950).
[130] B. G. Childs and A. D. LeClaire, Acta Metall. **2**, 718 (1954).
[131] T. S. Kê, J. Appl. Phys. **19**, 285 (1948).
[132] J. Wenig and E. S. Machlin, Acta Metallurgica **4**, 262 (1956).
[133] V. S. Postnikov, Dokl. Akad. Nauk SSSR **71**, 79 (1953).
[134] V. S. Postnikov, Труды Кемеровского горного института (Trans. Kemerovo Mining Inst.) **2**, 277 (1956).
[135] V. S. Postnikov and G. K. Mal'tseva, Труды Кемеровского педагогического института (Trans. Kemerovo Pedagogical Inst.) **3** (1957) (in press).
[136] C. L. Ang, J. Sivertman, and C. Wert, Acta Metallurgica **8**, 558 (1955).
[137] V. S. Postnikov, Thesis, Moscow Steel Inst., 1952.
[138] Pridantsev, Meshcherinov, and Piguzov, Dokl. Akad. Nauk S.S.S.R. **111**, 98 (1956).
[139] V. S. Postnikov, Труды Кемеровского

педагогического института (Trans. Kemerovo Pedagogical Inst.) **3** (1957) (in press).

[140] T. S. Kê and P. T. Jung, Scientia Sinica **3**, 261 (1955).

[141] T. S. Kê and C. M. Wang, Scientia Sinica **4**, 501 (1955).

[142] T. S. Kê and C. T. Tsien, Scientia Sinica **5**, 625 (1956).

[143] T. S. Kê and J. L. Ma, Scientia Sinica **6**, 81 (1957).

[144] C. T. Kung and T. S. Kê, Scientia Sinica **6**, 223 (1957).

[145] K. M. Rozin and B. N. Finkel'shtein, Dokl. Akad. Nauk SSSR **71**, 811 (1953).

[146] V. S. Postnikov and R. S. Lebedev, Труды Кемеровского педагогического института (Trans. Kemerovo Pedagogical Inst.) **3** (1957); Физика металлов и металловедение (Phys. of Metals and Metal Res.) **6** (1958) (in press).

[147] T. A. Read, Trans. AIME **143**, 30 (1941).

[148] S. G. Holder, E. E. Stansbury, and J. H. Frye, J. Metals **8**, 993 (1956).

[149] R. R. Hasiguti, Collection: Proc. Internat. Conf. Theor. Phys., Tokyo, 577, 1954.

[150] K. Matta and J. Barducci, Ric. Scient. **26**, 1160 (1956).

[151] K. M. Entwistle, J. Inst. Metals **21**, 249 (1954).

[152] W. R. Thomas and G. M. Leak, Nature **176**, 29 (1955).

[153] J. L. Snocu, Physica **7**, 711 (1941).

[154] C. Wert and C. Zener, Phys. Rev. **76**, 1169 (1949).

[155] J. D. Fast and L. J. Dijustra, Phil. Tech. Rev. **13**, 172 (1951).

[156] J. D. Fast and J. L. Meijering, Philips Research Rep. **8**, 1 (1953).

[157] L. J. Dijustra and R. J. Sladek, J. Metals **5**, 69 (1953).

[158] W. Köster and L. Bangert, Arch. Eisenhüttenw. **25**, 231 (1954).

[159] T. S. Kê, P. T. Jung, and J. N. Wang, Scientia Sinica **4**, 263 (1955).

[160] P. Stark, B. L. Averbach, and M. Cohen, Acta Metallurgica **4**, 91 (1956).

[161] T. S. Kê and J. L. Ma, Scientia Sinica **5**, 19 (1956).

[162] P. T. Jung and T. S. Kê, Scientia Sinica **5**, 645 (1956).

[163] B. N. Finkel'shtein and L. F. Usova, Труды МИС (Trans. Moscow Steel Institute) **36**, 176 (1957).

[164] J. Barducci and P. Gence, Ric. Scient. **26**, 2080 (1956).

[165] E. Scheil and W. Thiele, Arch. Eisenhüttenw., **12**, 103 (1938).

[166] Yu. V. Piguzov, Dokl. Akad. Nauk S.S.S.R. **112**, 636 (1957).

[167] F. A. Lewis, C. E. Roberts, and A. R. Ubbelohde, Proc. Roy. Soc., ser. A, **220**, 279 (1953).

[168] K. Bungardt and H. Preisendanz, Arch. Eisenhüttenw. **27**, 715 (1956).

[169] B. N. Finkel'shtein and N. S. Fastov, Dokl. Akad. Nauk S.S.S.R. **71**, 875 (1950).

[170] B. N. Finkel'shtein and N. S. Fastov, Проблема металловедения и физика металлов (Problems of Metal Research and the Physics of Metals) **2**, 245 (1951).

[171] N. S. Fastov, J. Exptl. Theoret. Phys. (U.S.S.R.) **22**, 487 (1952).

[172] V. S. Postnikov, Физика металлов и металловедение (Phys. of Metals and Metal Res.) **6**, 522 (1958).

[173] L. D. Landau and E. M. Lifshitz, Механика сплошных сред (Mechanics of Continuous Media) Gostekhizdat, 1953.

[174] C. Zener, Met. Techn. **13**, 1 (1946).

[175] R. Becker and M. Kornetsui, Z. Phys. **88**, 634 (1934).

[176] M. Kersten, Z. techn. Phys. **15**, 463 (1934).

[177] R. Becker and W. Döring, Ferromagnetismus, 1939.

[178] R. Bozorth, Ferromagnetism, N. Y. Van Nostrand, 1951.

[179] H. J. Williams and R. M. Bozorth, Phys. Rev. **59**, 939 (1941).

[180] J. C. Maxwell, Phill. Trans. Roy. Soc. **157**, 49 (1867); Phil. Mag. **35**, 129 (1868).

[181] Ya. I. Frenkel', Введение в теорию металлов (Introduction to the Theory of Metals) Gostekhizdat, 1950.

[182] A. S. Nowick, Phys. Rev. **82**, 340 (1951).

[183] V. S. Gorskii, Phys. Z. Sow. **8**, 443 (1935); J. Exptl. Theoret. Phys. (U.S.S.R.) **6**, 272 (1936).

[184] V. S. Gorskii, Phys. Z. Sow. **6**, 77 (1934).

[185] W. Köster, Z. Metallkunde **32**, 145 (1940).

[186] W. Köster, Z. Metallkunde **32**, 151 (1940).

[187] W. Köster and A. Schneider, Z. Metallkunde **32**, 156 (1940).

[188] L. Landau and G. Rumer, Phys. Z. Sow. **11**, 18 (1937).

[189] A. S. Kompaneets, Dokl. Akad. Nauk S.S.S.R. **14**, 267 (1937).

[190] A. Akhiezer, J. Exptl. Theoret. Phys. (U.S.S.R.) **8**, 1318, 1330 (1938).

[191] S. N. Tkachenko, J. Exptl. Theoret. Phys. (U.S.S.R.) **9**, 314 (1939).

[192] B. V. Paranjape, Proc. Phys. Soc. ser. A, **66**, 572 (1953).

[193] H. O. Kneser and K. Voelz, Ann. d. Phys. **11**, 377 (1953).

[194] N. S. Fastov, Проблемы металловедения и

физика металлов (Problems of Metal Research and the Physics of Metals) **4**, 388 (1955).
[195] L. Gurevich, Основы физической кинетики (Fundamentals of Physical Kinetics) Gostekhizdat, 1940.
[196] J. M. Burgers, First Report on Viscosity and Plasticity, Amsterdam, 1935, p. 19.
[197] A. S. Nowick, Symposium on the Plastic Deformation of Crystalline Solids. Pittsburgh, 1950, p. 155.
[198] A. S. Nowick, J. Appl. Phys. **25**, 1129 (1954).
[199] T. A. Read, Phys. Rev. **58**, 371 (1940).
[200] G. H. Faund, Trans. AIME **161**, 120 (1945).
[201] J. H. Swift and J. E. Richardson, J. Appl. Phys. **18**, 417 (1947).
[202] R. R. Hasiguti and T. Hirai, J. Appl. Phys. **22**, 1084 (1951).
[203] E. Körber and M. Hempel, Inst. Eisenforschung **22**, 1 (1935).
[204] L. Kukanov, Заводск. лабор. (Plant Laboratory) **6**, 722 (1937).
[205] W. Köster, Z. Metallkunde **32**, 282 (1940).
[206] T. Norton, Trans. AIME **137**, 49 (1940).
[207] F. Seitz and T. A. Read, J. Appl. Phys. **12**, 100 (1941).
[208] C. Zener, H. Clarke, and C. Smith, Trans. AIME **147**, 90 (1942).
[209] T. A. Read and E. P. T. Tyndall, J. Appl. Phys. **17**, 713 (1946).
[210] Ch. Boulanger, C. R. Acad. Sc. **226**, 1170 (1948).
[211] T. S. Kê, J. Metals **2**, 575 (1950).
[212] T. S. Kê, J. Metals **2**, 581 (1950).
[213] T. S. Kê and C. Zener, Chinese J. Phys. **8**, 133 (1951).
[214] J. Weertman and J. S. Kochler, J. Appl. Phys. **24**, 624 (1953).
[215] H. Fusfeld, J. Appl. Phys. **24**, 1062 (1953).
[216] W. Köster and E. Stolte, Z. Metallkunde **45**, 356 (1954).
[217] W. Köster, L. Bangert, and R. Hahn, Arch. Eisenhüttenw. **25**, 596 (1954).
[218] L. A. Glikman and E. A. Khein, J. Tech. Phys. (U.S.S.R.) **24**, 400 (1954).
[219] L. A. Glikman and E. A. Khein, J. Tech. Phys. (U.S.S.R.) **24**, 560 (1954).
[220] J. Friedel, Ch. Boulanger, C. Crussard, Acta Metallurgica **3**, 380 (1955).
[221] A. S. Nowick, Acta Metallurgica **3**, 312 (1955).
[222] J. Weertman, J. Appl. Phys. **26**, 202 (1955).
[223] V. S. Postnikov and M. M. Belyaev, Физика металлов и металловедение (Phys. of Metals and Metal Res.) **2**, 504 (1956).
[224] V. S. Postnikov, Труды Кемеровского горного института (Trans. Kemerovo Mining Inst.) **1**, 202 (1954).
[225] A. V. Grin', Физика металлов и металловедение (Phys. of Metals and Metal Res.) **4**, 383 (1957).
[226] T. A. Kontorova and Ia. I. Frenkel', J. Exptl. Theoret. Phys. (U.S.S.R.) **8**, 1340 (1938).
[227] J. D. Esheby, Proc. Roy. Soc., ser. A. **197**, 395 (1949).
[228] G. Leibfried, Z. Phys. **127**, 344 (1950).
[229] J. S. Kochler, Imperfections in Nearly Perfect Crystals, 1952, p. 197.
[230] P. G. Bordoni and M. Nuovo, Nuovo cimento **11**, 127 (1954).
[231] W. Köster and L. Bangert, Acta Metallurgica **3**, 274 (1955).
[232] A. Cranato and K. Lücke, J. Appl. Phys. **27**, 583 (1956).
[233] A. L. Nowick, J. Appl. Phys. **22**, 925 (1951).
[234] V. T. Shmatov, Report at the all-Union Inter-university Conference on Relaxation Phenomena in Pure Metals and Alloys, April 1958, Moscow.
[235] T. S. Kê, Phys. Rev. **76**, 549 (1949).
[236] V. S. Postnikov, Труды Кемеровского горного института (Publ. Kemerovo Mining Inst.) **1**, 202 (1954).
[237] C. Zener, Phys. Rev. **71**, 34 (1947).
[238] A. S. Nowick, Phys. Rev. **88**, 925 (1952).
[239] E. Woordruff; Phys. Rev. **16**, 321 (1903).
[240] C. Wert, J. Appl. Phys. **21**, 1196 (1950).
[241] T. S. Kê, P. W. Jang, Scientia Sinica **6**, 623 (1957).
[242] M. A. Isakovich, J. Exptl. Theoret. Phys. (U.S.S.R.) **18**, 386 (1948).
[243] J. Weertman, J. Appl. Phys. **28**, 193 (1957).
[244] A. H. Cottrell, Dislocations and Plastic Flow in Crystals, Oxford, 1953.
[245] G. Schoeck, Phys. Rev. **102**, 1458 (1956).
[246] M. Lessen, J. Appl. Phys. **28**, 364 (1957).
[247] I. I. Chernikova, Физика металлов и металловедение (Physics of Metals and Metal Research) **5**, 102, 176 (1957).
[248] B. N. Finkel'shtein and L. F. Usova, Сборник трудов Московского института стали (Collection of papers of the Moscow Steel Inst.) **36**, 176 (1957).
[249] Pavlov, Kriuchkov and Fedotov, Физика металлов и металловедение (Physics of Metals and Metal Research) **5**, 371 (1957).
[250] H. J. McSkimin, J. Appl. Phys. **24**, 988 (1953).
[251] H. E. Bömmel, Phys. Rev. **96**, 220 (1954).
[252] W. P. Mason and H. E. Bömmel, J. Acoust. Soc. **28**, 930 (1956).
[253] A. B. Pippard, Phil. Mag. **46**, 1104 (1955).

Translated by G. Volkoff

THE PRESENT STATE OF THE PROBLEM OF ACCELERATION OF ATOMIC PARTICLES *

V. I. VEKSLER

Usp. Fiz. Nauk 66, 99-110 (September 1958)

IN the course of the past few years experimental physics has made extraordinary advances in the problem of artificially obtaining particles with energies of hundreds of Mev. Only a few years ago, such particles could be observed only in cosmic rays, and then with negligible intensity.

The artificial production of beams of protons and electrons with enormous energy made it possible to investigate and discover a variety of phenomena of fundamental importance for nuclear physics and for all of science. A new and very promising field of nuclear physics has come into being—the physics of high-energy particles. The tempo of developments in this new field has been extraordinary. The production of intense beams of mesons shed new light on the nature of nuclear forces and the discovery of the existence of new elementary particles—heavy neutral mesons, antiprotons, and antineutrons. This whole stream of new facts is due to the development of methods for acceleration of charged particles. An independent branch of experimental physics, devoted to accelerators, has developed. This field depends on the latest achievements of radio engineering and is related most directly with an advanced state of the electronic and radio-engineering industry.

The importance of fast particles for the study of the properties of nuclei was first made clear by the celebrated experiments in which Rutherford succeeded in producing the disintegration of nitrogen nuclei by bombarding them with the α particles from the natural decay of RaC.

The rapid development of nuclear physics was accompanied by the production of artificial nuclear "artillery"—machines which can impart high energies to atomic particles, electrons, and protons.

The decisive step in this direction was made with the invention of the cyclotron by Lawrence, who first applied the resonance method to the acceleration of charged particles. As you know, nuclear physics owes many very important achievements to the cyclotron. However it already became apparent at the end of the Thirties that the solution of the problem of nuclear forces required the development of accelerators capable of producing beams of particles with energies far greater than those which could be obtained using the cyclotron. Ever since, physicists have tried to develop accelerators giving particles with greater and greater energies.

Allow me to give a brief explanation of the reasons for this situation.

In 1937 there were discovered in the cosmic rays charged particles with a mass intermediate between the masses of electron and proton. These particles were called mesons. Soon after, the existence of several types of mesons was established. This discovery was essentially the beginning of a new chapter in the development of our concepts of the nature of nuclear forces and the structure of nucleons. Physicists were faced with the problem of developing and inventing accelerators which could produce mesons artificially, and of using the mesons for the investigation of the nature of nuclear forces.

The study of cosmic rays had shown that the collision of high-energy nucleons with atomic nuclei is an efficient method for producing mesons. The mass-energy relation of the theory of relativity shows that, in order to produce a new particle by the collision between a particle "projectile" and a nucleon or nucleus at rest, the accelerated particle must be given an energy at least equal to (and, in fact, even greater than)

$$W = M_0 c^2,$$

where M_0 is the rest mass of the particle to be produced in the collision process, and c is the velocity of light.

The rest mass of the meson is several hundred times as great as that of the electron, and corresponds to an energy of approximately 150 Mev. At the time of the discovery of mesons, physicists had no means for artificially producing such particles. This problem could not be solved by using the cyclotron.

You know that the principle of operation of the cyclotron is the use of resonance between the frequency of revolution of the protons moving in the magnetic field of the cyclotron and the frequency of the alternating electric field that accelerates the protons.

Roughly speaking, such a resonance exists only so long as the velocity of the protons moving in the

*Report to the General Session, Academy of Sciences, U.S.S.R., March 27, 1958.

cyclotron is sufficiently smaller than the velocity of light. However as the energy rises the velocity of the particles increases and then, in accordance with the theory of relativity, the mass of the proton increases. As a result the matching of the period of revolution of the particles and the frequency of the accelerating electric field becomes poorer and poorer, until finally the cumulative detuning from resonance makes further acceleration impossible.

The maximum energy to which protons can be accelerated in the cyclotron is around 10 Mev and is thus one-twentieth the energy needed for the artificial production of mesons.

All sorts of attempts were made to overcome the difficulties related to the relativistic increase in mass of the accelerated particles. It appeared (and this opinion prevailed for many years) that the relativistic increase in mass is a fundamental difficulty and sets a limit on the possibility of accelerating protons to high energies.

In 1944 I and (somewhat later) MacMillan succeeded in showing that the relativistic mass increase not only is not a hindrance to the development of efficient methods of acceleration of charged particles but, on the contrary, the relativistic effect makes it possible to accomplish the resonant acceleration of electrons and protons to extremely high energies. The key to this advance was the discovery of the phenomenon of autophasing (phase focusing; Transl. note). Autophasing made possible the development of a whole variety of accelerators, both for the acceleration of electrons as well as for the acceleration of nuclear particles, protons, and deuterons.

I shall not enumerate all the types of accelerators based on the use of autophasing, but shall mention only the main ones.

The first of these is the cyclic electron accelerator called the synchrotron. One of the first machines of this type, with an energy of 30 Mev, was already built in 1947 in the P. N. Lebedev Physics Institute of the Academy of Sciences, U.S.S.R.

At present there are many different synchrotrons operating in various countries, giving beams of electrons with energies up to a few hundred Mev. Synchrotrons already exist giving electron beams with an energy of 1.5 Bev, and synchrotrons for 7 Bev are under construction. As you know, in the Soviet Union for many years the synchrotron of the Physics Institute has operated at 280 Mev and the synchrotron of the Leningrad Physico-Technical Institute at 100 Mev.

Synchrotrons enable us to obtain ultra-hard electromagnetic radiation. The existence of the neutral π meson was established using such machines. High-energy electrons and photons are among the most efficient tools for investigating the structure of nucleons, the fundamental building blocks from which atomic nuclei are made.

The next important type of accelerator is the phasotron (or synchrocyclotron). It is used to accelerate protrons, deuterons, and α particles, and is a very efficient generator of π mesons. There are at present more than ten phasotrons operating throughout the world. The largest of them is the one constructed in the Soviet Union under the direction of Meshcheryakov, D. V. Efremov and A. A. Mints. Almost all we know about the role of π mesons in the problem of nuclear forces has come from the use of powerful beams of mesons generated in phasotrons.

The third type of accelerator is the well known synchrophasotron (proton synchrotron; Transl. note). Machines of this type are the American cosmotron at 3 Bev, the bevatron at 6.3 Bev and, finally, the synchrophasotron constructed in our country, which gives a beam of 10-Bev protons. The 10-Bev synchrophasotron belongs to the Joint Institute for Nuclear Research. Its parameters have been reported many times in the press, and I shall not talk about them. Intensities of approximately 10^9 particles per pulse have been obtained with this accelerator, and physical investigations on rapidly-increasing have already begun. It is to accelerators of this type that we owe the vital knowledge we have concerning the physics of elementary particles. These machines opened the way to the discovery of long-lived neutral mesons, antiprotons, and antineutrons.

All the accelerators enumerated above are of the cyclic type. They are all characterized by the fact that the particles to be accelerated move in a closed trajectory and traverse the same electric field over and over again. There is still another type of resonance accelerator, the linear accelerator, in which the particles to be accelerated move in a straight line. The rapid development of these machines is also linked to the discovery of the phenomenon of autophasing.

I shall try to give a very general picture of the principle of autophasing, without going into details about the special features of the operation of this mechanism in each particular type of accelerator.

Every cyclic accelerator contains two fundamental elements:

1. A magnetic field that guarantees the cyclic character of the motion of the charged particles.

2. An accelerator structure, in which an alternating electric field is excited, and which is in-

tended to communicate energy to the charged particles. The frequency of the electric field can be either constant or variable in time. The magnetic field can also be constant or increase with time. These two elements are enough to cause the appearance of the mechanism of autophasing.

Allow me to use a few very simple formulas to make the operation of the autophasing mechanism clear.

Let us try to give an expression for the time spent by the particles moving in the magnetic field in going through one revolution. This time T will obviously be determined by the ratio of the path length s traversed by the particle during one revolution to the velocity of the particle v; i.e.,

$$T = \frac{s}{v} .$$

For charged particles moving in a magnetic field which is uniform and constant (or almost constant) in time, we know that the trajectory is a circle whose radius R is given by

$$R = \frac{Mvc}{He} = \frac{M_0 vc}{\sqrt{1-\frac{v^2}{c^2}}He} ,$$

where H is the magnetic field strength, $M = M_0/\sqrt{1-v^2/c^2}$ is the total mass of the particle, e its change, and v is its velocity.

Thus the path traversed by the particle during one revolution will be $s = 2\pi R = 2\pi Mcv/He$. Consequently, we get the time T for one revolution by dividing the path s by the velocity of the particle, i.e.,

$$T = \frac{2\pi M_0 c}{\sqrt{1-\frac{v^2}{c^2}}He} .$$

According to relativity theory, $Mc^2 = w$, so we get $\frac{2\pi}{He}\frac{w}{c} = \frac{2\pi}{ec}\frac{w}{H}$.

This formula contains essentially all that we need for understanding the mechanism of autophasing. It shows that in every cyclic accelerator there is a simple relation between the three fundamental quantities: the strength of the magnetic field which controls the motion of the particles, the period of revolution of the particles, and their energy. Consequently our problem is, using the relation between these quantities, to find and establish conditions in which the particle energy can be increased continually by making appropriate changes in one or both of the other parameters. It turns out that the solution of the problem is essentially very simple. We must use the already known resonance method of acceleration. But the amplitude of the potential difference of the electric field which accelerates the particles can no longer be chosen arbitrarily but must satisfy a simple condition.

The most significant feature of this method of acceleration is that, because of the dependence of the particle mass on velocity, any deviation of the time of revolution of the particles from the resonance value causes a change in the energy increment given to the particles by the electric field, so that the period of revolution automatically returns to its resonance value. It turns out that to produce such a mechanism we can vary the magnetic field or the frequency of the electric field slowly, or we can keep them both constant and apply sufficiently large potential differences to the accelerator structure. If for some reason the particle acquires energy too rapidly during acceleration, the resonance between its period of revolution and the frequency of the electric field is spoiled, and the particle begins to pick up less energy in each revolution and is then brought back into resonance. On the other hand, if for some reason the particle energy rises too slowly during the acceleration to maintain resonance, the particle begins to receive greater increments of energy from the electric field, and the necessary additional energy is thus automatically given to it.

Such an automatic maintenance of balance between the period of revolution of the particle and the period of the accelerating field can be achieved in cyclic accelerators of the most varied types. For example, one can keep the magnetic steering field constant in time and decrease the frequency of the accelerating electric field according to an arbitrary law. One can also do the reverse: increase the magnetic field and keep the frequency constant. Finally, one can change both. Phasotrons are accelerators of the first type, in which the magnetic field is constant, synchrotrons are of the second type, with constant frequency, and synchrophasotrons, which I have already mentioned, are instruments of the third type.

In all cases, in order to have autophasing it is sufficient to satisfy a very simple inequality relating the potential difference of the accelerating electric field to the rate of change of the magnetic field or the frequency.

Compared with the cyclotron, autophasing accelerators have made it possible to raise the limiting energy by a factor of 1000, and apparently this figure is not a final limit.

Permit me now to turn to another characteristic feature of the autophasing accelerator, which is very important for the consideration of the prospects of future development of charged-particle accelerators.

The energy increase achieved by using the prin-

ciple of autophasing is not gotten for nothing. For example, while the ordinary cyclotrons have pole diameters of 1 or 1.5 m and weigh from a few tens to hundreds of tons, the present-day phasotrons already weigh several thousand tons and their pole diameters are 5 to 7 meters. The weights of synchrophasotrons go up into the tens of thousands of tons, while the radius of the electromagnet is tens of meters, and in accelerators under construction even becomes hundreds of meters. This tendency of present day accelerators was cleverly underscored by the famous Italian physicist Fermi, who remarked jokingly in one of his last lectures that if one extrapolates the present ratio of maximum particle energy to accelerator dimensions, then to obtain particles with an energy of 10^{16} ev one would have to build an accelerator with an orbit diameter equal to the diameter of the Earth and with a vacuum chamber encircling the equator.

This is a facetious description. But it undoubtedly describes the main physical features of present accelerators. Furing the past 15 years physicists have worked to get higher and higher energies. Autophasing enabled them to solve this problem. But the consequence of autophasing is that the intensity, i.e., the flux of particles, which we can get from the accelerator always decreases when the particle energy increases. For example, one can obtain a current of around 100 milliamperes from the cyclotron, while the phasotron gives only a microampere. Synchrophasotrons like the bevatron or the accelerator at the Joint Institute give a current of only 10^{-3} microamperes. Lawrence pointed out that, if the situation develops further in this way, we might get in Fermi's fantastic accelerator an intensity of one proton per day.

Thus, together with the tremendous increase in energy there occurs a marked decrease in the intensity of the flux of accelerated particles. The point is that as the dimensions and weight of the accelerator increase, there is a very rapid rise in the power requirements of the electromagnets of the cyclic accelerators and of the high-frequency generators used to produce the accelerating electric field. Since the rating of the power sources is limited, the result is that the number of pulses of accelerated particles produced by the accelerator in one second decreases catastrophically. In the cyclotron the number of pulses is 10^7 per sec, in the phasotron of the Joint Institute this number is already only 100 per sec, while in the 10-Bev synchrophasotron we get a pulse every 12 seconds.

This situation is very typical, since it is intimately related to the acceleration principle which we are utilizing and which requires that the dimensions and weight of the accelerator must increase in order to raise the energy of the particles.

There are two classes of problems, two tremendous fields of research, which require an even greater increase in the energy of charged particles and, at the same time, a further increase in the intensity of the flux of accelerated particles. One of these fields is in the domain of basic research: the study of the nature of elementary particles.

The recent advances in physics of which I spoke earlier require the continual development of accelerators with higher and higher energies. To produce π mesons we must impart an energy of 150 Mev or more to the particle "projectiles", to observe antiprotons we need about 6 Bev, and to produce pairs of the so-called cascade hyperons we already require about 10 Bev. There can be no doubt that further discoveries will emphasize this tendency and in turn demand still greater energies from accelerators.

At the same time, there is a second and possibly no less important class of problems for the solution of which we need to develop powerful accelerators capable not only of producing very high energies of the order of tens of Bev or possibly even higher, but at the same time a very intense flux of accelerated particles.

I cannot discuss these problems in any detail. I mention only that they are related to the possibility of practical use of accelerators and may encompass the most varied aspects of everyday life.

I should like to try and discuss very briefly, almost sketchily, the outlook for developments in the two directions just mentioned. Let me begin with the question of maximum energy. At present the most energetic particles are gotten from cyclic accelerators. The maximum particle energy which can be obtained in a cyclic accelerator with magnetic field strength H is practically independent of the type of accelerator, and is determined only by the strength of the magnetic field which can be used to maintain the particles in their orbit, and by the radius of the orbit. In all presently known materials, the limiting field strength does not exceed 30,000 oersteds. Therefore, if we want to continue in the direction of increasing particle energies using existing methods, there is only one way available to us—that of simply increasing the dimensions of the accelerator. But here we shall rapidly get to the limit of what is reasonable, and we are already close to the limit. Earlier the physicist did not have to concern himself with problems of an economic nature. But in the de-

sign of more and more powerful accelerators, essentially only the technological and economic factors are now decisive. Everyone will understand that if the magnet of an accelerator weighs around 40,000 tons, if we need to erect enormous buildings and foundations for it, if we must build power stations giving a pulsed supply of hundreds of thousands of kilowatts, then unfortunately these problems cannot be avoided by physicists. But purely economic factors alone do not limit us in simply increasing the dimensions of existing accelerators. It can be shown that the engineering difficulties in producing tremendous electromagnets, their weight, the rigidity of the requirements for their magnetic characteristics, etc.— all these factors increase, as the dimensions of the electromagnet increase, roughly as the 2.5 power or cube of the limiting energy.

What is the reason for such a rapid increase in the dimensions of accelerators? It turns out that it is caused by a single physical phenomenon, the satisfying of the conditions for stability of the motion of the particles. In any cyclic accelerator, and especially in an accelerator which makes use of autophasing, the particles go through an enormous number of revolutions during the process of acceleration. For example, in the synchrophasotron of the Joint Institute, during the acceleration time, which lasts 3.3 seconds, the particles go through a few million revolutions and traverse a path of a million kilometers. If, over such a long path, random effects such as collisions with gas molecules, perturbations of the motion due to inhomogeneity of the magnetic field, etc, are not to result in a catastrophic loss of accelerated particles, it is necessary that the motion of the particles be stable. We know that in order to have stable conditions the magnetic field must have a very definite configuration that guarantees cyclic motion of the particles.

The theory shows that in all existing accelerators the magnetic forces which assure stability and keep the particles near the equilibrium trajectory are in general very small, so that the amplitudes of oscillation of the particles about the equilibrium trajectory are large. Thus the region of space in which the particles move, and in which the magnetic field configuration must satisfy the stability conditions, must also be made quite large. Increasing the energy of the accelerator and correspondingly increasing the orbit radius makes it necessary to increase the height and width of the region of space in which the particles move. This results in an increase of the weight of the electromagnet proportional to the cube of the orbit radius. In the synchrophasotron of the Joint Institute the width of the "track" in which the particles move is about one-and-a-half meters. If we wanted to increase the maximum particle energy to 30 or 50 Bev, using the same stability conditions, we would have to build an electromagnet weighing close to a million tons, which is obviously already completely unrealistic.

A few years ago a group of American physicists proposed a very ingenious idea for getting around this difficulty. The method developed by these physicists provided a new approach to the stability problem and has been called "strong focusing." It was shown that one can get a marked increase in the magnetic forces, which guarantee stability of the particle motion, by making the magnetic field configuration vary periodically with azimuth.

I cannot spend the time to discuss this question in any detail. I may only say that this new idea makes it possible to reduce the amplitude of oscillation of the particles about their equilibrium orbit by almost an order of magnitude, and thus to reduce correspondingly the weight of the electromagnet and the power required for it. This makes it possible for us to go further in the direction of obtaining still higher energies.

At present accelerators based on the principle of strong focusing are being designed and constructed all over the world. In the Soviet Union, under the direction of V. V. Vladimirskii, E. G. Komar, and Mints, with Efremov assisting, a 50-Bev strong-focusing accelerator is under design. The weight of the proposed accelerator will be about 30,000 tons. But the gap will be only 10 cm wide and 6 cm high, even though the orbit radius will be approximately 250 meters. Accelerators of a similar type, designed for 30 Bev, are being constructed at a rapid pace in Switzerland and America. In these accelerators, as in the usual synchrophasotrons, the acceleration is accomplished by using the autophasing principle. We may anticipate that the strong focusing method will enable us to go several times higher in energy than the figures already achieved at present.

I should say, however, that the advance made by using strong focusing is purchased at a very high price. I am thinking of the extremely difficult problem of eliminating resonant blowup of the particle oscillations. The theory shows that in such systems the requirements on the precision of all parts of the accelerator are raised by a tremendous amount.

To give a qualitative indication of the difficulties with which one has to cope, I mention that for a 50-Bev strong-focusing synchrophasotron,

having a ring magnet radius of approximately 250 meters, the settling of the foundation on which the magnet is to be placed must be uniform (strictly speaking, the 30'th harmonic must be kept uniform) to 0.1 mm, even though the total weight of the electromagnet is some tens of thousands of tons and the perimeter of the electromagnet is about 1.5 kilometers. Extremely rigid requirements must also be imposed on the constancy of the magnetic field configuration and on the precision of all sorts of other parts of this gigantic accelerator. As you see, even though these problems are apparently technically solvable, their solution is still an extremely difficult job.

For just these reasons, the strong-focusing method cannot enable us to make a marked advance upward in the energy scale. Despite all the ingenuity of this method, its use enables us only to raise the maximum particle energy but little. The problem of getting a large increase in intensity of particle beams and of moving on to ultrarelativistic particles with energies in the hundreds and thousands of Bev, which is important from many points of view, cannot be solved by the methods I have been describing. We must look for some completely new direction if we want to plan for a rapid advance.

I shall now turn to the last part of my talk, which will necessarily contain a relatively large element of speculation, and which concerns the prospects for accelerator development.

I shall first talk about an idea in which the two questions, the question of getting large currents and the question of getting high energy particles, are linked together in a surprising way.

Let us put this simple question to the experimental physicist: "What do you want to do with a high energy particle?" The experimenter gives the trivial answer that he uses the accelerated particle as a projectile which interacts with a particle at rest in a target, and observes scattered particles, meson production, antinucleons, etc.

But why should the target be at rest? There appears to be a possibility of observing processes at ultrahigh energies if the target (which until now was always at rest) itself moves with high velocity opposite to the beam of particles which we use as projectiles.

Let us see what happens if the projectile and target move opposite to one another with equal relativistic velocities. It is easy to show that if two protons having energies of say 10 Bev move opposite to one another and collide, the whole interaction process will proceed as if one of the protons was at rest while the other moved with an energy of approximately 20 Bev. Thus if we make two beams of high energy particles move opposite to one another and observe the processes which occur in the collision, we will be able to investigate phenomena which would occur if we had at our disposal an accelerator giving particles with an ultra-high energy equal to twice that of the individual particles in the colliding beams.

Such a method for setting up processes which could be produced only by particles with superhigh energies seems to me to be very promising. But it is not so easy to realize in practice. This is because the cross section for the collision of, say, two protons is a minute quantity. To observe such collisions it would thus be necessary, as a computation shows, to have in each of the colliding beams an enormous current, of the order of 50 to 100 amp. It is doubtful whether one can get such currents from existing autophasing accelerators. The maximum pulsed currents which can be attained at present are 0.1-0.01 amps. Let me make it clear that I am now speaking of instantaneous currents and not of the average current which I talked about earlier.

To make possible the practical realization of the colliding-beam method, we must raise the instantaneous currents by a factor of 500 to 1000. There are all sorts of ingenious ideas for doing this. In 1953, V. A. Petukhov, M. S. Rabinovich, and A. A. Kolomenskii in the Soviet Union, and somewhat later Kerst, Simon, et al. in the U. S. A., proposed new magnetic systems with a magnetic field which is constant in time and which would make possible a considerable increase in the current of accelerated particles.

There are suggestions for using existing accelerators for current storage. It is proposed to build magnetic storage systems, in which particle beams from present-day accelerators can be stored for a long time, after which by some means one brings about a collision between the beams accumulated in two storage vessels. We are of course a long way off from realizing these rather crude proposals; I mention them only to point out the somewhat unexpected connection between the two problems of obtaining high currents and of getting ultrahigh energies.

The principles of operation of all the accelerators which I have discussed up to now are essentially related to the problem of the motion of an isolated particle in given magnetic and electric fields. The rapid development of accelerators was made possible by the setting up of a rigorous theory which enabled accurate calculation of particle motions. The basic theory was

formulated by the Soviet theoretical physicists M. S. Rabinovich and A. A. Kolomenskii, and also by some physicists in other countries.

The new ideas which I am now going to discuss briefly belong to the realm of phenomena in which one must take account of the collective interaction of particles. For this reason a lot of the material is still qualitative. I am thinking of the relativistically stabilized beam proposed by A. M. Budker, the coherent method of acceleration proposed by me, and Ya. I. Feinberg's idea of plasma waveguides.

Although all three ideas are completely different, they have a common feature. All contemplate the use of a plasma for producing intense beams of high-energy particles.

For lack of time, in the preceding discussion I touched only on questions relating to the present state and prospects of development of the cyclic method of acceleration. In the course of the past ten years, under the influence of the stormy progress in radio technology and the use of the principle of autophasing, linear accelerators have also begun to develop rapidly. Despite the fact that they have all sorts of advantages for high energy work, linear accelerators are still not as efficient as cyclic accelerators. One of the main reasons for this is that in ordinary resonators and waveguides it is not possible to produce and make sufficiently efficient use of very high electric field strengths. Also, in these accelerators it is difficult to achieve simultaneously the conditions for phase and spatial focusing of particles. In 1956 Ya. I. Feinberg pointed out the possibility of simultaneously overcoming both these difficulties by using a plasma, at rest or in motion, placed in a longitudinal magnetic field. Such a plasma acts like a waveguide, and has all sorts of surprizing properties. Waves having a wavelength much greater than the transverse dimensions of the guide can propagate through such a plasma waveguide. The possibility arises of producing extremely high field intensities in just that small region of space in which the accelerated particles move. The difficulty in combining phase and radial stability of the accelerated particles disappears.

All these points promise to open new prospects for linear accelerators.

I shall now turn to the question of stabilized beams.

I mentioned earlier that in a cyclic accelerator the maximum energy that can be given to the particle is determined only by the magnetic field strength at the orbit.

It was first pointed out by A. M. Budker that, by using a relativistic plasma, one can establish a magnetic field approximately two orders of magnitude greater than can be gotten with any ferromagnet, i.e., one can obtain a field of order 10^6 oersteds. Budker's idea is based on the utilization of the relativistic properties of a stream of charged particles. He showed that if one produces a sufficiently high current in a cyclic accelerator by using relativistic electrons and compensates the repulsion of the electrons by means of positive ions, intense electromagnetic radiation will appear in the plasma. The appearance of radiation causes the diameter of the plasma tube to contract markedly, and a tremendous magnetic field develops inside the plasma and on its surface. One can make use of this magnetic field by placing particles inside the tube of plasma and then accelerating them by any standard method, such as autophasing. This idea is very beautiful, but its realization requires that many difficulties be overcome. Still, if it could be done we could count on getting particles with an energy of approximately 10^{11} ev.

In the method which I just described, the problem of producing very high magnetic fields is solved, while the acceleration of the particles is accomplished by the old method. A few years ago I pointed out the possibility of establishing a new principle of acceleration of atomic particles. It was called the coherent method. There appear to be all sorts of ways of realizing coherent acceleration. This principle can be used both for the acceleration of charged particles and for the acceleration of quasineutral aggregates. Allow me to give a very brief discussion of a few examples of this new mechanism.

Imagine that we have a small cluster of charges containing n positive ions. Suppose further that we send toward this cluster a stream of electrons which move with velocity v past the cluster. It is easy to show that then each particle in the cluster will experience an accelerating force, which is the greater the number of charges in the cluster. Thus the effective accelerating field can, at least in principle, be made very large, even reaching values of millions of electron volts per centimeter. A qualitative estimate shows that the efficiency of this mechanism is close to unity. Naturally this is a very attractive idea. But it must be said that there are great difficulties in the way of its realization, mainly related to the problem of obtaining high currents of relativistic electrons.

The next variant of coherent acceleration, which is of interest for many reasons, is the acceleration of quasi-neutral clusters. Imagine a quasi-

neutral cluster, consisting of electrons and positive ions or of electrons and positrons. We direct toward this cluster an electromagnetic wave whose wavelength is somewhat greater than the dimensions of the cluster. The electromagnetic wave will induce polarization oscillations of the electrons in the cluster. Part of the momentum of the electromagnetic wave will be scattered by these oscillations, as a result of which the whole cluster will pick up momentum and will begin to move in the direction of propagation of the wave. Obviously, this variant of coherent acceleration is essentially a special case (of course, under quite different conditions) of the phenomenon of light pressure discovered by P. N. Lebedev.

The most important advantage of this method is that in accelerating neutral globs of plasma one can accelerate more particles than in any other acceleration method. In principle, the coherent method gives one the hope of accelerating particles to energies of 10^{12} ev and even higher. It seems to us that the only way of obtaining such tremendous energies is by using impact acceleration, in which a cluster (or a current ring) of relativistic electrons collides with another cluster or current ring containing ions. It can be shown that if the mass of the relativistic cluster of electrons is much greater than the mass of the ion cluster, the ions will be given an enormous ultra-relativistic energy in the collision.

There are also other variants of coherent acceleration, but I shall not mention them here.

I was of course able to give only a very quick survey of these three methods, which are related to the use of plasmas and are still complete novelties for the physics of accelerators.

Recently the important role of plasma processes in nature has become clearer. A few years ago the possibility was pointed out of producing electromagnetic radiation by means of the Cerenkov effect. It was also pointed out that charged particles may be accelerated through the motion of streams of plasma in the magnetic fields of stars. Many attempts were made to relate the mechanism of cosmic-ray production to these processes. In these processes, stochastic or probability mechanisms of acceleration may play an essential role. We know that in the plasma of a gas discharge there occur very often fast particles whose energy is far greater than the potential difference applied to the plasma. One can show that the motion of clusters of plasma in the inhomogeneous magnetic fields of cosmic space must be accompanied by the production of relativistic electrons. So you see that in astrophysics the problem of the acceleration of charged particles is related most intimately to the peculiarities of plasma.

The struggle to attain those enormous energies in the millions of electron volts, which by some mechanism unknown to us are generated in cosmic space, the use of these artificial projectiles for the study of the nature of elementary particles—these are the most interesting problems confronting physicists and engineers working in this enchanting branch of science.

Translated by M. Hamermesh

NEW MAGNETIC MATERIALS—GARNET FERRITES

K. P. BELOV and M. A. ZAITSEVA

Usp. Fiz. Nauk **66**, 141-144 (September, 1958)

WITHIN the last year and a half or two years, great interest has developed in the study of a new class of ferrites that have the so-called garnet structure. This interest is caused by the fact that the garnet ferrites display properties favorable for their use in microwave applications. Their main advantage, as compared with ferrites with spinel structure, is an extremely narrow ferromagnetic resonance-absorption line. Crystals of some garnet ferrites give an absorption line-width of a few oersteds,[1,2] whereas crystals of spinel ferrites have a width of several tens of oersteds.

Apart from the practical aspect, the study of garnet ferrites is of great interest for the theory of magnetism and crystal chemistry, because their structure is quite different from the familiar structure of the spinel ferrites. Knowledge of their ferromagnetic properties gives additional information about the nature of spontaneous magnetization and its relation to the structure of the material.

The ferromagnetic properties of the garnet ferrites were discovered quite recently, in 1956, by the French scientists Bertaut and Forrat[3] and Pauthenet,[4] and independently a little later, in 1957, by the Americans Geller and Gilleo. Garnet ferrites have the formula

$$3\,Me_2O_3 \cdot 5\,Fe_2O_3 \text{ (or } Me_3Fe_2Fe_3O_{12}),$$

where Me is one of the rare-earth elements or yttrium.

The preparation of such a type of compound is carried out in accordance with usual ceramic practice, from iron oxides and oxides of rare-earth elements. Crystals are grown from the melt by slow cooling from 1325°C to about 900°C, with lead oxide as solvent.[6] The crystal lattice of a garnet ferrite is cubic; it contains a very large number of ions located at different crystallographic sites. Because of the complexity of the distribution of anions and cations within it, it is almost impossible to depict. The Fe^{3+} and Me^{3+} cations occupy spaces between the oxygen ions. These can be fourfold, the 24 d sites (tetrahedral spaces); sixfold, the 16 a sites (octahedral spaces); or eightfold, the 24 c sites. The Fe^{3+} ions are distributed over the d and a sites, the Me^{3+} ions over the c sites.

To explain the magnetic properties of the garnet ferrites, Néel[7] proposed to regard the lattice of these substances as composed of three sublattices (in contrast to the spinels, in which, of course, two sublattices are considered, with more octahedral than tetrahedral sites). The Fe^{3+} ions are distributed over two of these (the a and the d), with more tetrahedral than octahedral sites. The Me^{3+} ions are distributed over the third sublattice (the c). According to Néel, there is a strong negative interaction between the first two sublattices, a and d; as a result, there occurs a resultant spontaneous magnetization J (on account of the excess of "tetrahedral" magnetic "moments"). The Me^{3+} ions are magnetized by the weak field of sublattices d and a, and in consequence the moments of the c sites are directed opposite to the magnetization of sublattice d. The distribution of sublattice magnetizations can be pictured schematically as follows:

$$\underset{Me^{3+}}{\overset{c}{\longrightarrow}} \quad \underset{Fe^{3+}}{\overset{d}{\longleftarrow}} \quad \underset{Fe^{3+}}{\overset{a}{\longrightarrow}}.$$

This scheme provides a possibility of calculating the magnetic moments (per mole) of the garnet ferrites at 0°K. Thus for the case of gadolinium garnet, $3Gd_2O_3 \cdot 5Fe_2O_3$, we have (noticing that $M_c = 7\mu_B$ and $M_d = M_a = 5\mu_B$):

$$6M_c - (6M_d - 4M_a) = 6\times 7\mu_B - (6\times 5\mu_B - 4\times 5\mu_B) = 32\mu_B.$$

Experiment gives $30\mu_B$. In the case of yttrium garnet, since Y is nonmagnetic, we have $6M_d - 4M_a = 10\mu_B$, which also agrees with experiment. Thus analysis of the magnitudes of the magnetic moments in garnet ferrites corroborates Néel's "three-sublattice" hypothesis. Recently this hypothesis has been corroborated by neutron diffraction experiments.[8]

Thus in the case of the garnet ferrites we encounter a new form of uncompensated antiferromagnetism, caused by the nonequivalence of three magnetic sublattices. At the present time the following data are known on the magnetic properties of garnet ferrites. The temperature dependence of the spontaneous magnetization has been studied in the range 2.2 to 750°K, by the method of removal from the field, for the following garnet ferrites:

$3Me_2O_3 \cdot 5Fe_2O_3$, where Me = Gd, Tb, Dy, Ho, Er, Tu, Yb, Lu, Sm, Eu, or Y. For the majority of the ferrites, the dependence of magnetization on field satisfies the following relation:

$$\sigma_H = \sigma_s + \chi H,$$

where σ_S is the spontaneous magnetization, and χ is the susceptibility in strong fields and is independent of the field. The spontaneous magnetization σ_S was found by extrapolation to zero field of the straight part of the curve in strong fields.

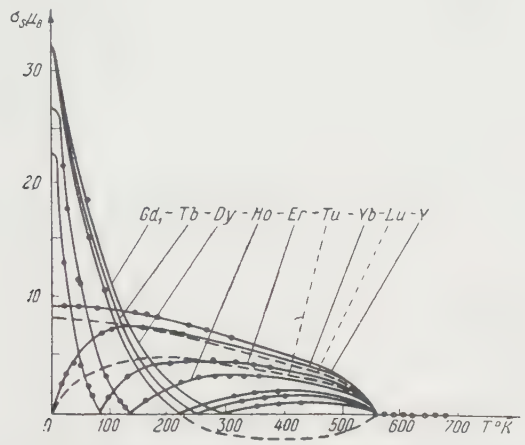


FIG. 1. Temperature dependence of the spontaneous magnetization σ_s (in μ_B per mol) for garnet ferrites.

In Figs. 1 and 2 are shown curves of the temperature dependence of the spontaneous magnetization (in μ_B per mol) for all the garnet ferrites investigated. In these figures there are two characteristic peculiarities that should be noted:

1. In the majority of the garnet ferrites (with the exception of the ferrites of Y, Lu, Sm, and Eu), there are observed compensation points, at which the magnetizations of the sublattices compensate each other.

2. It is a surprising and interesting fact that all these substances have about the same Curie point (in contrast to this, the molar magnetic moments at 0°K vary over a relatively wide range). The coincidence of the Curie points must be attributed to

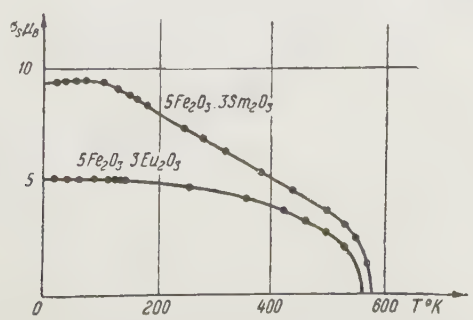


FIG. 2. Temperature dependence of the spontaneous magnetization σ_s (in μ_B per mol) for garnet ferrites of samarium and europium.

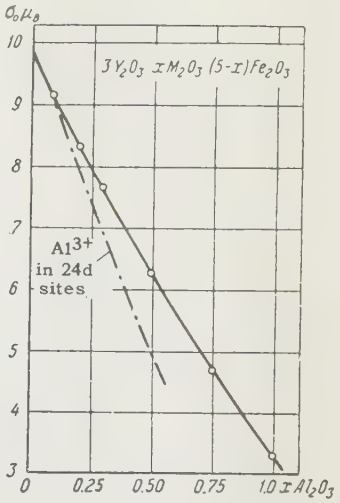


FIG. 3. Dependence of the spontaneous magnetization σ_0 (in μ_B per mol) on content of substitutional Al^{3+} ions.

the fact that the Me^{3+} ions take little part in the ferromagnetism; the interaction of the c lattice with the d and a is small, therefore here the magnitude of the Curie point is determined essentially by the strength of the interaction between the iron ions (the number of which remains approximately the same in all the garnet ferrites).

In the foreign literature there have recently[9-11] appeared numerous papers on the investigation of

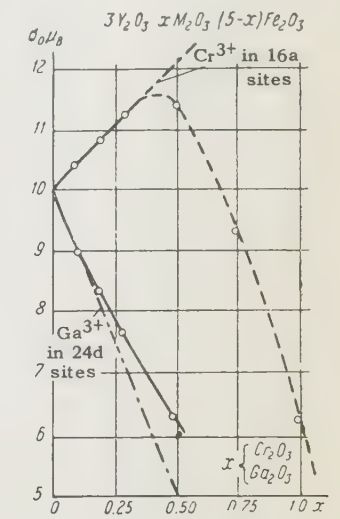


FIG. 4. Dependence of the spontaneous magnetization σ_0 (in μ_B per mol) on content of substitutional Cr^{3+} and Ga^{3+} ions.

the replacement of iron ions by other ions in the stoichiometric garnet ferrite $3Me_2O_3 \cdot 5Fe_2O_3$.

Substitutions of this type are directed at two goals: (1) investigation of garnet ferrites with new magnetic properties; (2) elucidation of the role of the sublattices in ferromagnetism; this can be accomplished by changing the number of magnetic ions in the lattices by introducing nonmagnetic or magnetic ions into them.

It should be noted also that experiments of this type are of interest for crystal chemistry. If ions with a different electron configuration and atomic radius are introduced into the sublattices, these

Magnetic moment $n\mu_B$ and Curie temperature θ for $3Y_2O_3 \cdot xMe_2O_3 (5-x) Fe_2O_3$

$x\,Me_2O_3$	$n\mu_B$ observed	θ, °K observed
0	9.92	545
0.25 Ga_2O_3	7.95	519
0.75 Ga_2O_3	4.35	460
0.34 Al_2O_3	7.00	497
1.00 Al_2O_3	3.25	415
0.25 Sc_2O_3	11.97	500
0.75 Sc_2O_3	14.4	365
0.50 In_2O_3	13.8	444
0.25 Cr_2O_3	10.95	515

ions will prefer fourfold, sixfold, or eightfold sites, and consequently it is possible thus to change the distribution of atoms in the garnet structure and to study their role in magnetization. Most of the experiments have been done by replacement of the iron in yttrium garnet ferrite by the ions: Al^{3+}, Ga^{3+}, Cr^{3+}, Sc^{3+}, In^{3+}. Introduction of Al^{3+} and Ga^{3+} ions decreases the magnetic moments of the garnet ferrites (cf. Figs. 3 and 4 and the table). This must be explained by supposing that Al^{3+} and Ga^{3+} are distributed on tetragonal sites (in sublattice d), in consequence of which the resultant magnetic moment decreases. Al^{3+} and Ga^{3+} have smaller ionic radii than Fe^{3+}, therefore they have an affinity for the tetrahedral positions. Sc^{3+} and In^{3+}, having larger radii than Fe^{3+}, preferentially occupy octahedral positions, therefore the magnetic moment of the ferrites increases (cf. table). The Cr^{3+} ion, though also smaller than the Fe^{3+}, nevertheless occupies octahedral positions, perhaps by virtue of its peculiar electronic configuration. This leads, in a certain Cr^{3+} concentration range, to an increase of magnetic moment.

Thus it follows from the experiments on substitution that the moment of a garnet ferrite can be increased by replacing an iron ion in sublattice a by any nonmagnetic ion, and can be decreased by replacing an iron ion in sublattice d by a nonmagnetic ion.

It should be noted that in all cases of replacement of an iron ion by another ion, the magnitude of the Curie temperature of the garnet ferrite decreases. This is explained as a consequence of the fact that the number of magnetic Fe^{3+}-O^{2-}-Fe^{3+} interactions on a magnetic ion in the garnet ferrite decreases. Summarizing everything that has been said about the magnetic properties of garnet ferrites, we must point out that we still have only scanty data on the magnetic properties of these interesting materials. Up to now nothing is known about hysteretic, magnetostrictive, electrical, and other properties of the garnet ferrites. Meanwhile these data are necessary for an understanding of the nature of their spontaneous magnetization.

[1] J. F. Dillon, Phys. Rev. **105**, 759 (1957); Calhoun, Overmeyer, and Smith, Phys. Rev. **107**, 993 (1957).

[2] J. Paulevé, Compt. rend. **244**, 1908 (1957); **245**, 1604 (1957).

[3] F. Bertaut and F. Forrat, Compt. rend. **242**, 382 (1956).

[4] R. Pauthenet, Compt. rend. **242**, 1859 (1956); Aléonard, Barbier, and Pauthenet, Compt. rend. **242**, 2531 (1956); R. Pauthenet, Compt. rend. **243**, 1499 and 1737 (1956); Néel, Bertaut, Forrat, and Pauthenet, Izv. Akad. Nauk SSSR, Ser. Fiz., **21**, 904 (1957) [Columbia Tech. Transl. **21**, 903 (1957)].

[5] S. Geller and M. A. Gilleo, Acta Cryst. **10**, 239 (1957).

[6] J. W. Nielsen, J. Appl. Phys. **29**, 390 (1958).

[7] L. Néel, Compt. rend. **239**, 8 (1954).

[8] Bertaut, Forrat, Herpin, and Mériel, J. phys. et radium **18**, 32s (1957).

[9] G. Villers and J. Loriers, Compt. rend. **245**, 2033 (1957).

[10] R. Pauthenet, J. Appl. Phys. **29**, 253 (1958).

[11] M. A. Gilleo and S. Geller, J. Appl. Phys. **29**, 380 (1958).

Translated by W. F. Brown, Jr.

RADIO ASTRONOMICAL INVESTIGATIONS WITH THE AID OF ARTIFICIAL SATELLITES

G. G. GETMANTSEV, V. L. GINZBURG, and I. S. SHKLOVSKII

Usp. Fiz. Nauk **66**, 157-161 (October, 1958)

THE use of satellites for optical astronomy is known to promise a possible increase in the spectrum of the received waves. Specifically, a satellite can receive the near and far ultraviolet, x-ray, and far infrared rays absorbed by the atmosphere before they reach the earth's surface. The situation is analogous in radio astronomy, since the range of wavelengths employed is also limited on one end by absorption in the troposphere and is limited on the other, long-wave end by refraction and absorption in the ionosphere.

Molecular (tropospheric) absorption begins to manifest itself at wavelengths shorter than 2 cm, and is particularly substantial at wavelengths of approximately 1.3 cm (water vapor) and approximately 5 mm (molecular oxygen). Consequently, reception of extraterrestrial millimeter waves without the special complications connected with the need for allowing for substantial absorption is possible only in the transparency region i.e., at wavelengths of approximately 8 mm (for more details on absorption in the atmosphere see reference 1, sections 84 and 120; radio astronomical measurements in the millimeter range are discussed in references 2, 3, and 4). An example of the difficulties encountered when going to shorter waves is that at wavelengths from 1 to 33 mm absorption in the atmosphere reduces the effective temperature of the solar radio waves by a factor of approximately 15.

Disregarding the use of very large mirrors, millimeter radiation can be received at the present time, even in the absence of absorption, only from the sun (effective temperature $T_{eff} \approx 6{,}000°$) and from the moon ($T_{eff} \approx 200°K$). Although an investigation of such radio waves is indeed of interest, it can hardly be considered as particularly important to the physics of the sun and the moon. On the other hand, the performance of such measurements with a satellite is difficult for many reasons (complexity of apparatus, need for orienting the receiving mirror on the sun or on the moon with an accuracy of several angular minutes).

The ionosphere begins to affect noticeably the reception of extraterrestrial radio waves at wavelengths on the order of 30 meters. It is enough to state that in daylight the critical wavelengths for the F_2 layer reach even 20 meters, meaning that longer waves, with the exception of some feeble effects, cannot reach the earth at all. At night the critical wavelength increases by a factor of several times. In addition, owing to the presence of a latitude dependence and to the seasonal and 11-year variations in the critical frequencies, it is possible to find conditions under which the ionosphere transmits waves up to 200 or 300 meters. Finally, the ionosphere sometimes lets longer waves pass, but these are generally greatly attenuated. Another factor complicating the earthside reception of long-wave cosmic radio waves is the high level of noise of terrestrial origin. By virtue of all these complications, the information available on cosmic radio waves in the region with $\lambda > 30$ m is quite meager.[5,6] This pertains in particular to wavelengths with $\lambda \gtrsim 100$ m, at which only a few individual measurements have been made;[7] the data obtained permit merely to state that there is strong radiation up to $\lambda \cong 600$ m, but the intensity of this radiation has not yet been established. Yet an investigation of cosmic radio waves with $\lambda > 30$ m is of primary importance. The point is that the radio waves represent magnetic bremsstrahlung of cosmic (relativistic) electrons moving in interstellar magnetic fields of intensity $\sim 10^{-5}$ oersteds. The frequency spectrum of the radio waves is determined here by the energy spectrum of the radiating particles. Consequently radio-astronomical measurements provide valuable data on the electronic component of cosmic rays in the galaxy (for more details see references 8 and 9). We note in particular that it is indeed in the wavelength region with $\lambda > 30$ m that a kink[5,6] appears in the spectrum of the cosmic radio waves. This circumstance, which requires further verification and refinement, is quite important since the location of the kink, which is evidence of the need for allowing for ionization losses, can be used to estimate the density of the interstellar gas.[8]

There is no doubt therefore that an investigation of the long-wave cosmic radio waves is in order. The use of artificial satellites for this purpose can solve the problem, since the interfering effect of

the ionosphere is substantially removed and the necessary apparatus is quite simple.*

To estimate the requirements that must be met by the receiving apparatus we recognize that in the wavelength region 10 cm < λ < 10 m the effective temperature of the non-thermal radiation is $T_{eff} = a\lambda^{2.8}$ (the intensity, $I_\nu = (2kT_{eff})/\lambda^2$, is proportional to $\lambda^{0.8}$). At a wavelength of approximately 3 m, T_{eff} is on the order of several thousand degrees. According to references 5 and 6, $I_\nu \cong$ const and $T_{eff} \sim 10^6$ or 10^7 at 30 m < λ < 100 m. At even longer wavelengths, λ > 100 m, no quantitative determination was made of the intensity of the radio waves that reach the top of the ionosphere. However, there are grounds for assuming[6] that $T_{eff} \gtrsim 10^7$ deg at λ > 100 m. At the same time at wavelengths greater than approximately 100 meters, the absorption in the interstellar medium already becomes substantial, and consequently upon further increase in λ the effective temperature of the cosmic radio waves will approach the temperature of the interstellar or interplanetary matter.

What apparatus should be used for the reception of such radio waves? Modern superheterodyne receivers with rf amplifier stages with extremely low noise factors, $S_n \sim 2$, can be built for wavelengths longer than 100 meters. The internal noise should originate mostly in the input circuits of the receiver. If the antenna is ideally matched to the receiver the ratio of the useful received signal to the internal noise should be T_a/T_{in}, where T_a is the antenna temperature, determined by the intensity of the received radio waves,† and T_{in} the temperature of the receiver input circuits. Naturally, it is impossible to obtain perfect matching between the antenna and the receiver. For a receiver designed for wavelengths longer than 100 meters, and for an antenna with a radiation resistances on the order of several tens of ohms, signal power gain on the order of 0.1 can be readily obtained.

Thus, the real signal-to-noise ratio at the grid of the first tube of the receiver should be on the order of $0.1\,T_a/T_{in}$. Assuming $F_nT_{in} \cong 10^3$ and $T_a \cong 10^7$, we have $0.1\,T_a/T_{in} \cong 10^3$, i.e., the internal noise is much smaller than the received signal. Under such conditions the relative error $\Delta T_a/T_a$ in the measurement of the intensity of a continuous-spectrum signal is determined, in the case of a sufficiently good receiver, only by the receiver bandwidth, Δf, and the averaging "time constant" of the receiver input circuits, τ. Here

$$\frac{\Delta T_a}{T_a} \sim \frac{1}{2}\frac{1}{\Delta f\tau}.$$

Putting, somewhat arbitrarily, Δf = 5 kcs and τ = 1 second, we get $\Delta T_a/T_a \cong 2\%$, i.e., the relative measurement error is small.

If the antenna-to-receiver signal transfer is not too small the radiation resistance R_Σ of the antenna, as already noted, must be on the order of 20 or 30 ohms. If a wire antenna (elementary dipole) is used we get, with $R_\Sigma = 80\pi^2(l/\lambda)^2 =$ 20 ohms, $l/\lambda \gtrsim 1/6$, where l is the length of the dipole. Thus, for example, $l \gtrsim 50$ m if λ = 300 m.

It is not convenient to use a wire antenna several tens of meters long on a satellite. A loop antenna with a ferrite core is structurally more suitable. Such an antenna has a relatively larger radiation resistance with small size and weight (length ~ 10 cm, weight ~ 300 g). The axis of the loop must be parallel to the metallic surface of the satellite. Naturally, the receiver must be calibrated after installation before launching.

Speaking of the installation of the apparatus on the satellite, the weight becomes an important parameter. A five-tube miniaturized superheterodyne receiver with the necessary gain at wavelengths greater than 100 meters and with provision for band switching (for example, tuned to several fixed frequencies) can be made relatively small. If one is interested only in the intensity of the cosmic radio waves, the apparatus must operate only for several revolutions of the satellite around the earth. The weight of the power supply will then be relatively small.

Since the satellite will have a rotation of its own and the orientation of the loop antenna will vary, the intensity of the received cosmic radio waves may vary somewhat. Therefore, generally speaking, it is necessary to know the orientation of the antenna loop at every instant of time. However, inasmuch as the cosmic radio waves are approximately isotropic in the range of interest to us, we can restrict use the maximum readings of the output meter. An important advantage of the observation of cosmic radio waves longer than 100 m from an artificial satellite is the great reduction in the atmospheric in other noise, since the receiver on the satellite is so to speak shielded

*We have in mind here the so-called "general" galactic radio waves. Long-wave radio waves from individual discrete sources can be observed only with very large antennas or interference apparatus. We consider here only the reception of "general" radiation.

†For equilibrium radiation with a temperature T_{eff}, the antenna temperature is naturally $T_a = T_{eff}$. This conclusion is true for any medium with a real dielectric constant ε > 0. In the ionosphere $T_a \sim T_{eff}$; a more exact connection between these quantities depends on many conditions ($T_a \cong T_{eff}$ above the F layer for waves reflected from this layer at any angle; for waves passing through the layer, $T_a \cong \frac{1}{2} T_{eff}$).

from the earth's surface when the satellite moves above the maximum of the F_2 layer.*

In conclusion we remark that data obtained with a satellite concerning the frequency spectrum of intensity of cosmic radio waves with wavelengths $\lambda \gtrsim 100$ m lead to certain conclusions regarding the electron concentration at the corresponding levels of the ionosphere above the maximum of the F_2 layer, not accessible to radio sounding from the surface of the earth. In fact, the antenna cannot receive waves for which $\epsilon(f, N) \cong 0$ at the point of reception. Neglecting the effect of the earth's magnetic field, we have

$$\varepsilon(f) = 1 - \frac{4\pi e^2 N}{m(2\pi f)^2} = 1 - 8 \cdot 10^7 \frac{N}{f^2},$$

where N is the electron concentration and f is the frequency of the received radiation. Even at very high altitudes, in interplanetary space, $N \sim 1$ to 5×10^2, i.e., $\epsilon(f) > 0$ when $f > f_0 = 9 \times 10^4$ to 2×10^5, or $c/f < \lambda_0 = 1.5$ to 3 km; when $N \sim 10^4$ the radiation propagates only if its wavelength in vacuum is $\lambda < \lambda_0 = 300$ m.

By measuring the frequency f_0 for which $\epsilon(f_0) = 0$ and reception ceases, it is possible to determine the concentration N.* The effect of the earth's magnetic field complicates the picture. However, for sufficiently high satellites this effect is small (owing to the reduced intensity of the earth's field). It can be accounted for even for lower satellites. For sufficiently long waves, longer than 200 to 300 meters, the influence of the field causes only one of the two normal waves propagating in the magnetoactive medium to be received. In this connection it is advisable to install two antennas and thus be able to measure the polarization.

*On the other hand, one could receive on the satellite not only cosmic radio waves but also waves due to the motion of fast electrons (of cosmic or solar origin) in the earth's magnetic field.[10]

An estimate given in reference 10, which can be readily refined and generalized, indicates that the presently known currents of fast electrons near the earth should produce magnetic bremsstrahlung radio waves that are considerably weaker than the cosmic radio waves, at least for wavelengths less than several kilometers. However, such long waves are particularly affected by the surrounding medium, i.e., by the magnetoactive plasma (in other words, it is necessary to reckon with the fact that the index of refraction of the medium differs from unity). This may give rise, in particular, to Cerenkov radiation. Such radiation, which is due to streams of particles of solar origin and produces auroras, has already been discussed in reference 11. It is very important that Cerenkov radiation, unlike magnetic bremsstrahlung radiation, is generated with equal efficiency by electrons and protons of a given velocity. E. A. Benediktov has called attention to the fact that the use of a satellite for the study of such radio waves may be very fruitful.

[1] Al'pert, Ginzburg, and Feinberg, Распространение радиоволн (Propagation of Radio Waves), Gostekhizdat, Moscow, 1953.

[2] J. P. Hagen, Astrophys. J. **113**, 547 (1951).

[3] Gordy, Ditto, Wyman, and Anderson, Phys. Rev. **99**, 1905 (1955).

[4] Whitehurst, Copeland, and Mitchell, J. App. Phys. **28**, 295 (1957).

[5] G. Reber and G. R. Ellis, J. Geophys. Res. **61**, 1 (1956).

[6] G. R. Ellis, Symposium on Radioastronomy, Melbourne, 1957.

[7] G. Reber, Symposium on Radioastronomy, Melbourne, 1957. J. Geophys. Res. **63**, 109 (1958).

[8] V. L. Ginzburg, Usp. Fiz. Nauk **62**, 37 (1957).

[9] I. S. Shkolovskii, Космическое радиоизлучение (Cosmic Radio Waves), Moscow, 1956.

[10] V. L. Ginzburg, Dokl. Akad. Nauk SSSR **76**, 377 (1951).

[11] G. R. Ellis, J. Atmos. Terr. Phys. **10**, 302 (1957).

Translated by J. G. Adashko

*The index of refraction of the ionosphere near the satellite influences, naturally, the input resistance of the antenna borne by it. To match the input resistance of the antenna to the receiver during the measurement of the radio waves, it is necessary to vary automatically the input resistance of the receiver by changing, for example, the capacitance of the tank circuit. On the other hand, if the satellite is provided with an instrument that measures the input impedance of the antenna, it is possible to measure directly the index of refraction in that region of the ionosphere where the satellite is located.

THEORY OF THE FERMI FLUID

(The Properties of Liquid He^3 at Low Temperatures)

A. A. ABRIKOSOV AND I. M. KHALATNIKOV

Usp. Fiz. Nauk **56**, 177-212 (October, 1958)

INTRODUCTION

THE study of the superfluidity of helium II gave rise to the question of the properties of quantum fluids, that is, of systems of interacting particles in whose behavior quantum effects play a major part. The theory of superfluidity given by L. D. Landau was at the same time the first theory of a quantum fluid. It described the properties of so-called Bose fluids, i.e., quantum systems of interacting particles for which the excitations obey Bose statistics.

Besides such fluids, however, there also exist others, forming a much more numerous class; these are the so-called Fermi fluids, whose excitations have spin 1/2 and obey Fermi statistics. These include liquid helium 3, the electrons in metals, and possibly heavy nuclei, although to be sure this last case is a quite special one. Until very recently there was no theory of the Fermi fluid. The theoretical calculations were usually confined to an approach analogous to that used for Bose fluids; some sort of energy spectrum was postulated, and this was then substituted into the Fermi distribution formula.

In view of the fact that in metals the situation was decidedly complicated by the anisotropy of the spectra, it was hard to test the correctness of such an approach. As for liquid helium 3 at low temperatures, the experimental data made it very clear that in this case such an approach could not be completely successful.

A consistent theory of Fermi fluids was developed by L. D. Landau in 1956.[1] He showed that the situtation is very different from that in Bose fluids, since in Fermi fluids a very important part is played by the interaction of the excitations, so that the excitations in a Fermi fluid cannot be regarded in all cases as an ideal gas.

In the present article we expound the theory of the Fermi fluid in its application to the isotropic model and show how the Landau theory provides an explanation of the various properties of liquid helium 3. In an appendix we present a recently completed research of L. D. Landau, which shows how the basic propositions of the theory of the Fermi liquid follow from a microscopic treatment of the interaction, and we also present there the theory of a rarefied Fermi gas. In this case one can obtain the basic quantities of the theory of the Fermi fluid by an actual calculation.

Before passing to the exposition of the theory, we recall for convenience in later applications, the main properties of liquid helium 3. The helium isotope of atomic weight 3 liquefies at 3.2°K (critical point $T_c = 3.3°$, $p_c = 845$ mm Hg), and remains liquid down to the lowest temperatures at which it has been studied (~ 0.2°K) and at pressures up to 30 atmos. The explanation is the same as in the case of helium 4: the weakness of the interaction between the atoms, and the smallness of the atomic mass. Owing to the latter circumstance the de-Broglie wavelength corresponding to the motion of helium 3 atoms at low temperatures can be larger than the distances between atoms; that is, the liquid becomes a quantum fluid.

Down to the lowest temperatures that have been used liquid helium 3 does not become a superfluid. Since according to the Landau theory every Bose liquid must possess superfluidity, while a Fermi liquid does not have this property, helium 3 must be regarded as a Fermi fluid. It must be stated that this fact is not trivial and does not follow automatically from the fact that helium 3 atoms have spin 1/2. A system of such atoms might have exclusively Bose excitations. Moreover, as we shall see later, such excitations actually exist, although they do not play any important part in determining the properties of helium 3. It is only the absence of superfluidity that enables us to conclude with assurance that helium 3 is a genuine Fermi fluid.

1. THE ENERGY OF THE EXCITATIONS

The possibility of describing an excited state of liquid helium 3 by means of a gas of quasi-particles and a corresponding distribution function in the energy scale is based on the fact that the interaction of the particles of a Fermi gas decreases rapidly as the temperature is lowered. In fact, it

is clear from general considerations that the time between collisions is proportional to the square of the diffuseness of the Fermi distribution, that is, to the mean square of the temperature. At the same time the mean energy of the Fermi particles in the diffuse part of the distribution (measured from the limit energy) is itself proportional to the first power of the temperature, and it is these particles that determine the macroscopic properties of the system. Owing to this, the "energy uncertainty" arising from the collisions will, at low enough temperatures, be small in comparison with the mean energy itself. It can also be said that at sufficiently low temperatures the damping of the excitations, described by an imaginary term added to their energy, will be small in comparison with the excitation energy.

Two fundamental assumptions lie at the foundation of the Landau theory. The first is that the classification of the levels of the Fermi fluid corresponds to the classification of the levels of noninteracting atoms. This means that as one gradually turns on the interaction the atoms gradually go over into "quasi-particles", each of which has a definite energy. Thus the number of quasi-particles is equal to the number of atoms.

The second assumption is that the interaction of the quasi-particles can be taken into account by means of a self-consistent field of the surrounding particles, which manifests itself in the fact that the energy of the system is not the sum of the energies of the quasi-particles, but instead is a functional of their distribution function. The energy of the quasi-particles must be defined as the variational derivation of the energy density in momentum space, i.e.,

$$\delta E = \int \varepsilon \delta n \, d\tau, \tag{1.1}$$

where $d\tau = 2dp_x dp_y dp_z/(2\pi h)^3$. The factor 2 in $d\tau$ arises from the fact that the quasi-particles have spin 1/2.

In some cases it is necessary to include a spin dependence of the particle energy. Owing to the fact that the spin is a quantum-mechanical quantity, we must regard the distribution function in this case as a statistical operator, and replace (1.1) by a definition of ϵ in the form

$$\delta E = \frac{1}{2} \mathrm{Sp}_\sigma \int \varepsilon \, \delta n \, d\tau. \tag{1.2}$$

The definition of the particle energy by Eq. (1.2) has the consequence that the equilibrium distribution function is in fact the Fermi function. To prove this it is most convenient to use the expression for the entropy*

$$S = -\frac{1}{2} \mathrm{Sp}_\sigma \int \{n \ln n + (1-n) \ln (1-n)\} \, d\tau, \tag{1.3}$$

since this formula is of purely combinatorial origin. From the condition that the entropy be a maximum for constant number of particles and constant energy,

$$N = \frac{1}{2} \mathrm{Sp}_\sigma \int n \, d\tau = \text{const}, \quad E = \text{const}$$

we can find the distribution function by taking the variation with respect to n:

$$n(\varepsilon) = n_F(\varepsilon) = \frac{1}{e^{\frac{\varepsilon - \mu}{T}} + 1}. \tag{1.4}$$

The energy ϵ, being a functional of n, depends on the temperature. This dependence can be put in the following form. If we denote by $\epsilon_0(\mathbf{p}, \sigma)$ the equilibrium energy of the quasi-particles for T = 0, then for a small departure from equilibrium or for small values of T it will be given by the formula

$$\varepsilon = \varepsilon_0(\mathbf{p}, \sigma) + \delta\varepsilon(\mathbf{p}, \sigma) = \varepsilon_0(\mathbf{p}, \sigma) + \frac{1}{2} \mathrm{Sp}_{\sigma'} \int f(\mathbf{p}, \sigma; \mathbf{p}', \sigma') \, \delta n(\mathbf{p}', \sigma') \, d\tau'. \tag{1.5}$$

Here $\delta n = n - n_F$ (T = 0), and f is an operator depending on the momenta and spin operators of two particles. In view of its definition, as the second variational derivative of E with respect to δ n, the function f must be symmetric under interchange of $\mathbf{p}, \sigma$ with $\mathbf{p}', \sigma'$. The function f is a very important quantity characterizing the Fermi fluid. As was shown by Landau, it is related to the forward scattering amplitude of two quasi-particles (cf. Appendix 2).

2. THE EFFECTIVE MASS

Because we are concerned here with very low temperatures, the energy ϵ_0 in the absence of an external field can be written in the form

$$\varepsilon_0 - \mu(0) = v(p - p_0), \tag{2.1}$$

where p_0 is the Fermi limiting momentum and v is the velocity at the Fermi limit. This velocity can be written in the form

$$v = \frac{p_0}{m^*}, \tag{2.2}$$

where m* is the effective mass. As has been shown by Landau, there is a definite relation between m*

*Here and throughout we use energy units for the temperature, i.e., k = 1.

and f, which is a consequence of the Galilean principle of relativity.

According to this principle the momentum per unit volume must be equal to the mass flux of the fluid. Because the velocity for the quasi-particles can be defined as the derivative $\partial\epsilon/\partial\mathbf{p}$, and the number of quasi-particles is equal to the number of atoms, we get:

$$\frac{1}{2}\mathrm{Sp}_\sigma \int \mathbf{p} n\, d\tau = \frac{1}{2}\mathrm{Sp}_\sigma \int m \frac{\partial \varepsilon}{\partial \mathbf{p}} n\, d\tau, \qquad (2.3)$$

where m is the mass of an atom.

Let us take the variation of this relation with respect to n:

$$\frac{1}{2}\mathrm{Sp}_\sigma \int \mathbf{p}\delta n\, d\tau = \frac{1}{2}\mathrm{Sp}_\sigma \int m \frac{\partial \varepsilon}{\partial \mathbf{p}} \delta n\, d\tau$$

$$+ \frac{1}{2}\mathrm{Sp}_\sigma \mathrm{Sp}_{\sigma'} \int\int mn \frac{\partial}{\partial \mathbf{p}} f(\mathbf{p}, \sigma, \mathbf{p}', \sigma')\, \delta n'\, d\tau\, d\tau'.$$

Interchanging $\mathbf{p}, \sigma$ and $\mathbf{p}', \sigma'$ in the last term and using the fact that δn is arbitrary, we get:

$$\frac{\mathbf{p}}{m} = \frac{\partial \varepsilon}{\partial \mathbf{p}} + \frac{1}{2}\mathrm{Sp}_{\sigma'} \int \frac{\partial f}{\partial \mathbf{p}'} n'\, d\tau' = \frac{\partial \varepsilon}{\partial \mathbf{p}} - \frac{1}{2}\mathrm{Sp}_{\sigma'} \int f \frac{\partial n'}{\partial \mathbf{p}'} d\tau. \qquad (2.4)$$

If the system is not in an external magnetic field, the relation (2.4) does not lose its generality when we take its trace with respect to the unprimed spin variable $(1/2\,\mathrm{Sp}_\sigma)$. In the case of equilibrium at $T = 0$, ϵ in Eq. (2.4) can be taken in the form (2.1), and $\partial n/\partial p$ can be replaced by $-\delta(p - p_0)$. We then find:

$$\frac{1}{m} = \frac{1}{m^*} + \frac{p_0}{2(2\pi\hbar)^3}\mathrm{Sp}_\sigma \mathrm{Sp}_{\sigma'} \int f(\chi) \cos\chi\, d\Omega, \qquad (2.5)$$

where $f(\chi)$ is the value of the function f for $|\mathbf{p}| = |\mathbf{p}'| = p_0$. Naturally, apart from a small correction, this relation also holds for temperatures near $T = 0$ and for small departures from equilibrium.

3. THE HEAT CAPACITY AND ENTROPY

Here we shall assume no external magnetic field, so that the quantities will not depend on the spin. A knowledge of the energy spectrum makes it possible to determine the heat capacity of the Fermi fluid.[2] It can be found in the usual way, by differentiating the energy with respect to the temperature for a prescribed number of particles.

An important point here is that the change of the energy density is given by Eq. (1.1); owing to this we can write

$$C = \left(\frac{\partial E}{\partial T}\right)_N = \int \varepsilon \left(\frac{\partial n}{\partial T}\right)_N d\tau \cong \int \varepsilon_0 \left(\frac{\partial n}{\partial T}\right)_N d\tau. \qquad (3.1)$$

In the present case the replacement of ϵ by ϵ_0 is justified by the fact that whereas the main term in the heat capacity is linear in T the correction in Eq. (1.5) gives only cubic terms.

In fact, the δn in Eq. (1.5) can be put in the form

$$\delta n = n(T) - n(0) = \int_0^T \frac{\partial n}{\partial T}\, dT$$

$$= \int_0^T \frac{\partial n}{\partial \varepsilon}\left(-\frac{\varepsilon - \mu}{T} + \frac{\partial \varepsilon}{\partial T} - \frac{\partial \mu}{\partial T}\right) dT.$$

As is well known, at low temperatures the derivative of the Fermi function has the form

$$\frac{\partial n}{\partial \varepsilon} \approx -\delta(\varepsilon - \mu) - \frac{\pi^2}{6} T^2 \frac{\partial^2 \delta(\varepsilon - \mu)}{\partial \varepsilon^2}.$$

In view of this we have

$$\delta\varepsilon = \int f \delta n'\, d\tau' = \frac{\pi^2}{6} T^2 \left[\frac{\partial}{\partial \varepsilon'}\left(f \frac{d\tau'}{d\varepsilon'}\right)\right]_{\varepsilon'=\mu}$$

$$- \int_0^T \left[f\left(\frac{\partial \varepsilon}{\partial T} - \frac{\partial \mu}{\partial T}\right)\frac{d\tau'}{d\varepsilon'}\right]_{\varepsilon',\ \mu} dT. \qquad (3.2)$$

The derivative $\partial\mu/\partial T$ is equal to the entropy, which, as we shall see, is linear in T. As for the term in $\partial\epsilon'/\partial T$, its order of magnitude can be established by Eq. (3.2). Differentiating this formula with respect to T, we find without difficulty that $\partial\epsilon/\partial T$ is of the first order in T, and consequently $\delta\epsilon$ is of the second order.

Because of the replacement ϵ by ϵ_0 the energy can be calculated simply from the integral

$$E = \int \varepsilon_0 n\, d\tau \quad \text{with} \quad N = \int n\, d\tau = \text{const},$$

where, again to within an error of cubic terms in T, we can replace ϵ by ϵ_0 in the functions n appearing in the integrals for E and N. After this the calculation of the integrals does not differ from that for the usual case of a Fermi gas.

In this way we find:

$$C = \frac{C}{N} = \gamma T; \quad \gamma = \frac{\pi^2}{3N}\left(\frac{d\varepsilon}{dp}\right)_{p=p_0} = \left(\frac{\pi}{3N}\right)^{\frac{2}{3}} \frac{m^* p_0}{\hbar^2}. \qquad (3.3)$$

Because of the linear dependence on the temperature the entropy is equal to the heat capacity.

By comparing Eq. (3.3) with the experimental data on the entropy of liquid He^3 at low temperatures we can find the parameter γ, and consequently can also determine the effective mass. Unfortunately, at present the experimental curve has only been found as far as the beginning of the linear part.[3] From these data we can obtain only an approximate value of γ, which turns out to be about 3 cal mole^{-1} deg^{-2}. Because of the equality of the number of

quasi-particles to the number of particles, we can find the Fermi limit momentum from the density of liquid He^3 ($\rho = 0.078$ g/cm^3). In this way we get:

$$m^* = 1.43 m_{He^3}, \quad \frac{p_0}{\hbar} = 0.76 \cdot 10^8 \text{ cm}^{-1}. \qquad (3.4)$$

4. THE MAGNETIC SUSCEPTIBILITY

Because particles possessing spin necessarily have a magnetic moment, the Fermi fluid must be magnetic. In the isotropic case there are only two possibilities. The fluid can be paramagnetic or ferromagnetic. Antiferromagnetism in an isotropic Fermi fluid is impossible. We present here only the case of the paramagnetic Fermi liquid, which corresponds to liquid He^3. A calculation of the susceptibility of such a liquid was given by Landau in reference 1 (cf. also reference 2).

When the system is in a magnetic field H, the particles naturally have an additional energy depending on their spins. If they were free particles, the change of the energy would be $-\beta(\sigma \cdot H)$. In the Fermi fluid, however, we have to take into account also the fact that there is a change in the distribution function. Thus we get

$$\delta\varepsilon = -\beta(\sigma \cdot H) + \frac{1}{2}\mathrm{Sp}_{\sigma'} \int f \delta n' d\tau'. \qquad (4.1)$$

In the present case the spin dependence of the function f is important. This dependence can be written in the general case in the form

$$f(\mathbf{p}, \sigma;\ \mathbf{p}', \sigma') = f(\mathbf{p}, \mathbf{p}') + \zeta_{ik}(\mathbf{p}, \mathbf{p}')\sigma_i \sigma_k'. \qquad (4.2)$$

It is natural to assume, however, that the spin interaction is mainly of exchange origin. In this case f takes the form

$$f(\mathbf{p}, \sigma;\ \mathbf{p}', \sigma') = f(\mathbf{p}, \mathbf{p}') + \zeta(\mathbf{p}, \mathbf{p}')\sigma \cdot \sigma'. \qquad (4.3)$$

The change of the distribution function appearing in Eq. (4.1), δn, is due only to the change of the energy ϵ, since the chemical potential changes only in the second order in H. Because of this we can set $\delta n = \partial n/\partial \epsilon\ \delta\epsilon$ in Eq. (4.1). This gives

$$\delta\varepsilon = -\beta(\sigma \cdot H) + \frac{1}{2}\mathrm{Sp}_{\sigma'} \int f \frac{\partial n'}{\partial \varepsilon'} \delta\varepsilon' d\tau'. \qquad (4.4)$$

This is the equation for $\delta\epsilon$, which can be assumed to have the form

$$\delta\varepsilon = -\xi(\sigma \cdot H). \qquad (4.5)$$

Substituting this into Eq. (4.4), we get the equation

$$\xi = \beta + \frac{1}{4} \int \zeta \frac{\partial n'}{\partial \varepsilon'} \xi' d\tau'. \qquad (4.6)$$

In view of the fact that $\partial n'/\partial \epsilon' \cong -\delta(\epsilon' - \mu)$, the integration is taken over the surface of the Fermi sphere. We shall see below that only the value of ξ taken on the surface of the Fermi sphere is of importance for the magnetic susceptibility. In this case both the arguments of $\zeta(p, p')$ have the absolute value p_0, and ζ depends only on the angle between them. Denoting $\int \zeta(\theta)\, d\Omega/4\pi$ by $\bar{\zeta}$, we get from Eq. (4.6):

$$\xi = \frac{\beta}{1 + \frac{1}{4}\bar{\zeta}\left(\frac{d\tau}{d\varepsilon}\right)_\mu}. \qquad (4.7)$$

The susceptibility is found from the formula

$$\chi = \frac{\partial}{\partial \mathbf{H}} \frac{1}{2} \mathrm{Sp} \int n_i^3 \sigma\, d\tau.$$

The value of this expression is of course determined by just the spin-dependent change of n. Thus we find:

$$\chi = \frac{1}{4}\beta\xi\left(\frac{d\tau}{d\varepsilon}\right)_\mu = \frac{\frac{1}{4}\beta^2\left(\frac{d\tau}{d\varepsilon}\right)_\mu}{1 + \frac{1}{4}\bar{\zeta}\left(\frac{d\tau}{d\varepsilon}\right)_\mu}. \qquad (4.8)$$

The quantity $(\partial\tau/\partial\epsilon)_\mu$ can be expressed in terms of the coefficient γ in the linear law of the heat capacity; this gives

$$\frac{1}{\chi} = \beta^{-2}\left(\bar{\zeta} + \frac{4\pi^2}{3\gamma N}\right). \qquad (4.9)$$

Equation (4.9) can be compared with the experimental data on the magnetic susceptibility of He^3;[4] this shows that $\bar{\zeta}$ is negative and is of magnitude 0.85 times the second term. Thus the exchange interaction has a decided effect on the magnitude of the magnetic susceptibility of liquid He^3. The sign of the effect is such that the exchange interaction facilitates parallel orientation of the spins. It does not lead to ferromagnetism, however, since the Fermi tendency toward an antiparallel arrangement of the spins prevails. It is possible that at higher pressures the parameters change in such a way that ferromagnetism becomes possible. There are no experimental data up to the present, however, to support such a conclusion.

5. THE KINETIC EQUATION

The kinetic equation for the distribution function in the absence of a magnetic field has the usual form

$$\frac{\partial n}{\partial t} + \frac{\partial n}{\partial \mathbf{r}} \cdot \frac{\partial \varepsilon}{\partial \mathbf{p}} - \frac{\partial n}{\partial \mathbf{p}} \cdot \frac{\partial \varepsilon}{\partial \mathbf{r}} = I(n), \qquad (5.1)$$

where I(n) is the collision integral. Here, however, we have to take into account the fact that the energy ϵ is a functional of the distribution function, and thus also depends on the coordinates. By means of this kinetic equation Landau[1] obtained expressions for the fluxes of energy and momentum.

To find the momentum flux one must multiply (5.1) by p_i and integrate over the phase space. Conservation of momentum in the collisions makes the integral $\int p_i I(n) d\tau$ vanish, and we have the equation

$$\frac{\partial}{\partial t}\int p_i n\, d\tau + \int p_i\left(\frac{\partial n}{\partial x_k}\cdot\frac{\partial \varepsilon}{\partial p_k} - \frac{\partial n}{\partial p_k}\cdot\frac{\partial \varepsilon}{\partial x_k}\right) d\tau = 0.$$

This expression in parentheses in the second integral can be written

$$\frac{\partial}{\partial x_k}\left(n\frac{\partial \varepsilon}{\partial p_k}\right) - \frac{\partial}{\partial p_k}\left(n\frac{\partial \varepsilon}{\partial x_k}\right).$$

Substituting this into the integral and integrating the second term by parts, we get:

$$\frac{\partial}{\partial x_k}\int p_i \frac{\partial \varepsilon}{\partial p_k} n\, d\tau + \int n\frac{\partial \varepsilon}{\partial x_i} d\tau.$$

The last term can be put in the form

$$\frac{\partial}{\partial x_i}\int n\varepsilon\, d\tau - \int \varepsilon\frac{\partial n}{\partial x_i} d\tau.$$

But in view of the fact that $\delta E = \int \epsilon \delta n d\tau$, the last term here is equal to $\partial E/\partial x_i$. Thus we get the equation

$$\frac{\partial}{\partial t}\int p_i n\, d\tau + \frac{\partial \Pi_{ik}}{\partial x_k} = 0. \qquad (5.2)$$

This equation expresses the law of conservation of momentum. The tensor Π_{ik} is the flux of momentum. It is given by

$$\Pi_{ik} = \int p_i \frac{\partial \varepsilon}{\partial p_k} n\, d\tau + \delta_{ik}\left[\mathrm{Sp}\int n\varepsilon\, d\tau - E\right]. \qquad (5.3)$$

In a similar way we also get the law of conservation of energy. To do this we multiply the kinetic equation by ϵ and integrate with respect to $d\tau$. Because of the conservation of energy in the collisions the integral $\int \epsilon I(n) d\tau$ vanishes, and we get:

$$\int \varepsilon\frac{\partial n}{\partial t} d\tau + \int \varepsilon\left(\frac{\partial n}{\partial \mathbf{r}}\cdot\frac{\partial \varepsilon}{\partial \mathbf{p}} - \frac{\partial n}{\partial \mathbf{p}}\cdot\frac{\partial \varepsilon}{\partial \mathbf{r}}\right) d\tau = 0.$$

Writing the term in $\partial n/\partial \mathbf{r}$ in the form

$$\frac{\partial}{\partial \mathbf{r}}\int n\varepsilon\frac{\partial \varepsilon}{\partial \mathbf{p}} d\tau - \int n\frac{\partial \varepsilon}{\partial \mathbf{r}}\frac{\partial \varepsilon}{\partial \mathbf{p}} d\tau - \int n\frac{\partial^2 \varepsilon}{\partial \mathbf{r}\,\partial \mathbf{p}}\varepsilon\, d\tau$$

and integrating by parts with respect to $\mathbf{p}$ in the term in $\partial n/\partial \mathbf{p}$, we get the law of conservation of energy

$$\frac{\partial E}{\partial t} = \int \varepsilon\frac{\partial n}{\partial t} d\tau = -\operatorname{div}\mathbf{Q}, \qquad (5.4)$$

where the flux of energy is given by

$$\mathbf{Q} = \int n\varepsilon\frac{\partial \varepsilon}{\partial \mathbf{p}} d\tau. \qquad (5.5)$$

6. THE VISCOSITY

From (5.1) and the expressions (5.2) for the momentum flux and (5.5) for the energy flux we can determine the coefficients of viscosity and thermal conductivity of the Fermi fluid.[5]

We begin by finding the viscosity. To do this, suppose that there occurs a motion in the liquid with a certain velocity $\mathbf{u}$ which is a slowly varying function of position. In this case the distribution function will differ only slightly from the equilibrium value:

$$n = n_0 + \delta n, \qquad (6.1)$$

where

$$n_0 = \frac{1}{e^{\frac{\varepsilon_0 - \mathbf{p}\cdot\mathbf{u} - \mu}{T}} + 1}; \quad |\delta n| \ll n_0. \qquad (6.2)$$

The quantity δn is found from the kinetic equation (5.1). As usual, we must insert the function n_0 in the left member of the kinetic equation. Furthermore we shall assume that at the point considered $\mathbf{u} = 0$. Substituting Eq. (6.2) into (5.1) we find:

$$\frac{\partial n_0}{\partial t} - \frac{1}{3}\frac{\partial n_0}{\partial \varepsilon}\mathbf{p}\frac{\partial \varepsilon}{\partial \mathbf{p}}\operatorname{div}\mathbf{u} - \frac{1}{2}\frac{\partial n_0}{\partial \varepsilon}\left(p_i\frac{\partial \varepsilon}{\partial p_k} - \frac{1}{3}p_l\frac{\partial \varepsilon}{\partial p_l}\delta_{ik}\right) \times\left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} - \frac{2}{3}\delta_{ik}\frac{\partial u_l}{\partial x_l}\right) = I(n). \qquad (6.3)$$

We now transform $\partial n_0/\partial t$ and show that this expression also does not depend on the term in f in Eq. (1.5). According to Eq. (6.2) we can write

$$\delta n_0 = \frac{\partial n_0}{\partial \varepsilon_0}\left(\delta\varepsilon_0 - \delta\mu - \frac{\varepsilon_0 - \mu}{T}\delta T - \mathbf{p}\cdot\delta\mathbf{u}\right). \qquad (6.4)$$

Since the derivative $\partial n_0/\partial \epsilon_0$ in Eq. (6.4) is different from zero in a small neighborhood of the point $\epsilon_0 = \mu$, and is a rapidly changing function in this region, we can suppose the quantities in the parentheses evaluated at that point (the error will be of the relative order $(T/\mu)^2$). The variations $\delta\epsilon_0$ and $\delta\mu$ are arbitrary and are by no means equal to each other. On the other hand, the distribution function is normalized by the relation

$$\int n_0 d\tau = N,$$

where N is the number of atoms in unit volume. Taking the variation of this relation, we find:

$$\delta N = \int \frac{\partial n_0}{\partial \varepsilon_0}\left(\delta\varepsilon_0 - \delta\mu - \frac{\varepsilon_0 - \mu}{T}\delta T - \mathbf{p}\cdot\delta\mathbf{u}\right) d\tau \approx -\left[(\delta\varepsilon_0 - \delta\mu - \mathbf{p}\cdot\delta\mathbf{u})\frac{d\tau}{d\varepsilon_0}\right]_{\varepsilon_0=\mu}, \qquad (6.5)$$

since $\int (\partial n_0/\partial \epsilon_0)\, d\epsilon_0 = -1$. Comparing (6.5) and (6.4), we get:

$$\frac{\partial n_0}{\partial t} = -\frac{\partial N}{\partial t}\cdot\frac{\partial n_0}{\partial \varepsilon}\cdot\left(\frac{d\varepsilon_0}{d\tau}\right)_{\varepsilon_0=\mu}.$$

The number of atoms N satisfies the equation of continuity

$$\frac{\partial N}{\partial t} + N\operatorname{div}\mathbf{u} = 0. \qquad (6.6)$$

Thus the term in $\partial n_0/\partial t$ gives an addition to the term in div u, which as a result now takes the form

$$-\frac{\partial n_0}{\partial \varepsilon}\left(\frac{1}{3}\,p\,\frac{\partial \varepsilon_0}{\partial p}-N\left(\frac{d\varepsilon_0}{d\tau}\right)_{\varepsilon_0=\mu}\right)\operatorname{div} u. \qquad (6.7)$$

Using Eq. (2.1) and the equality of the number of excitations to the number of particles, which makes $p_0/h=(3\pi^2N)^{1/3}$, we find that the expression just written vanishes. Thus apart from small corrections of the order $(T/\mu)^2$ there are no terms in div u in the left member of the kinetic equation.

The right member of Eq. (5.1) is the collision integral

$$I(n)=-\int w\,[n_1n_2(1-n_1')(1-n_2')-(1-n_1)(1-n_2)n_1'n_2']$$
$$\times\delta(p_1+p_2-p_1'-p_2')\,\delta(\varepsilon_1+\varepsilon_2-\varepsilon_1'-\varepsilon_2')\,d\tau_2d\tau_1'dp_2'. \qquad (6.8)$$

Here we must keep in mind that the δ function of the energies has as its arguments the actual energies, i. e., the energies as given by Eq. (1.5). Owing to this, substitution of the function $n_0(\epsilon_0)$ into the collision integral does not make it vanish. The integral vanishes only when we substitute in it the true equilibrium function, i. e., $n_0(\epsilon)$. The total distribution function can be written

$$n=n_0(\varepsilon)-\frac{\partial n_0}{\partial \varepsilon}\int f\,\delta n\,d\tau+\delta n. \qquad (6.9)$$

We introduce the notation

$$\delta n=-\frac{n_0(1-n_0)}{T}\,\nu=\frac{\partial n_0}{\partial \varepsilon_0}\,\nu. \qquad (6.10)$$

Substituting this in Eq. (6.9), we find that the term added to the equilibrium distribution function also has the form (6.10). Instead of ν however, it contains the quantity

$$\psi=\nu+\left(\frac{d\tau}{d\varepsilon_0}\right)_\mu\int f(\chi)\,\nu\,\frac{d\Omega}{4\pi}. \qquad (6.11)$$

Substitution of Eq. (6.9) in the collision integral, with neglect of terms quadratic in ψ, gives:

$$I(n)=\frac{1}{T}\int w n_{01}n_{02}(1-n_{01}')(1-n_{02}')(\psi_1+\psi_2-\psi_1'-\psi_2')$$
$$\times\delta(p_1+p_2-p_1'-p_2')\,\delta(\varepsilon_1+\varepsilon_2-\varepsilon_1'-\varepsilon_2')\,d\tau_2d\tau_1'dp_2'. \qquad (6.12)$$

In general the collision probability w depends on all four momenta. The momenta of importance here, however, are those with magnitudes close to the limit value on the Fermi surface. We can therefore assume that w depends only on θ, the angle between p_1 and p_2, and φ, the angle between the planes (p_1, p_2) and (p_1', p_2').

Let us make use of the fact that the momenta of the particles in the main region of the integration differ little in magnitude from the limit momentum p_0. If we rotate the plane of the vectors (p_1', p_2')

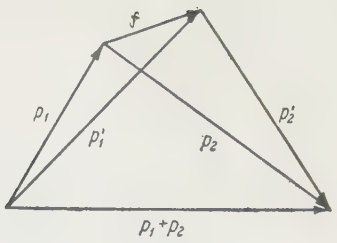


FIG. 1

through the angle φ relative to the axis in the direction p_1+p_2, so that this plane coincides with the plane (p_1, p_2), we get the diagram shown in Fig. 1. It is clear that the vector f will be small in magnitude and that the angles between all the momenta and the axis p_1+p_2 will be approximately equal to $\theta/2$ or $-\theta/2$. Thus we get

$$p_1'\approx p_1+f_z\cos\frac{\theta}{2}+f_r\sin\frac{\theta}{2},\quad p_2'\approx p_2-f_z\cos\frac{\theta}{2}+f_r\sin\frac{\theta}{2}, \qquad (6.13)$$

where f_Z is the component of f along the axis p_1+p_2 and f_r is the perpendicular component. We get rid of the δ function by integrating over dp_2', and replace the integral over $d\tau_1'$ by that over $df_rdf_Zd\varphi$, introducing a system of cylindrical coordinates with axis along p_1+p_2:

$$d\tau_1'=2p_0\sin\frac{\theta}{2}\,df_r df_z d\varphi\,\frac{1}{(2\pi\hbar)^3}. \qquad (6.14)$$

We now introduce the following notation:

$$x=\frac{\varepsilon_1'-\mu}{T},\quad t=\frac{\varepsilon_1-\mu}{T},\quad \varkappa=\frac{\mu}{T},\quad y=\frac{\varepsilon_2'-\mu}{T}; \qquad (6.15)$$

with the values of ϵ_i from Eq. (2.1). The change from the variables f_r and f_Z to x and y is accomplished very easily by means of Eq. (6.13). Then we get:

$$d\tau_1'\int d\tau_2\,\delta(\varepsilon)=\frac{m^{*3}T^2dx\,dy\left(\frac{d\Omega}{2\pi}\right)\left(\frac{d\varphi_2}{2\pi}\right)}{8\pi^4\hbar^6\cos\left(\frac{\theta}{2}\right)}, \qquad (6.16)$$

where we have denoted the angular differential $\sin\theta d\theta d\varphi$ by $d\Omega$, which involves the arguments of $w(\theta, \varphi)$. We note that because of the indistinguishibility of the particles the angle φ varies only from zero to π.

From considerations of symmetry it is clear that ψ must have the form

$$\psi=\frac{1}{2}q(t)\left(p_i\frac{\partial\varepsilon}{\partial p_k}-\frac{1}{3}p_i\frac{\partial\varepsilon}{\partial p_l}\delta_{ik}\right)\cdot\left(\frac{\partial u_i}{\partial x_k}+\frac{\partial u_k}{\partial x_i}-\frac{2}{3}\delta_{ik}\frac{\partial u_m}{\partial x_m}\right). \qquad (6.17)$$

When this expression is substituted in the collision integral the second factor can be transformed by means of the addition theorem for spherical harmonics. After the integration over the angle φ_2 only the first terms remain, i.e.,

$$P_2(\theta_1+\theta_2)\to P_2(\theta_1)P_2(\theta_2),\text{ where }P_2(\theta)=\frac{3}{2}\cos^2\theta-1.$$

As the result we get from Eq. (6.3) the following equation for q:

$$n_0(t)[1-n_0(t)] = \frac{m^{*3}T^2}{8\pi^4\hbar^6\cos\frac{\theta}{2}} \int \frac{d\Omega}{2\pi} \int_{-\infty}^{\infty} dx \int_{\substack{-x \\ x+y>t-x}}^{\infty} dy\, w(\theta,\varphi) \times n_0(t)\, n_0(x+y-t)[1-n_0(x)][1-n_0(y)] \times [q(t)+q(x+y-t)P_2(\theta)-q(x)P_2(\theta')-q(y)P_2(\theta_2')]. \quad (6.18)$$

In view of the fact that in the range of temperatures we are considering only values of x, y, and t much smaller than κ are of importance, we can take the lower limit of the integration in Eq. (6.12) to be $-\infty$. Under these conditions, assuming that q is a symmetric function (as is confirmed by the results) we easily verify that all the terms with different q's can be reduced to the same form as regards the dependence on x and y. Thus the bracketed expression takes the simple form

$$q(t)+q(x)[P_2(\theta)-P_2(\theta_1')-P_2(\theta_2')].$$

The change from the angles θ_1' and θ_2' to the angles θ and φ is easily obtained from Fig. 1. Integrating over the variable y, on which q no longer depends, after some manipulation we get the following equation:

$$\frac{8\pi^4\hbar^6}{m^{*3}T^2} = \overline{\frac{w(\vartheta,\varphi)}{\cos\frac{\theta}{2}}\cdot\left[\frac{3}{4}(1-\cos\theta)^2\sin^2\varphi-1\right]} \times\left\{\int_0^{\infty}\frac{dx\,x[q(x+t)+q(x-t)]}{e^x-1}+2\int_0^{\infty}\frac{dx\,xq(x)}{e^x+1}+\int_0^{t}dx\cdot xq(x-t)\right\} + \overline{\frac{w(\theta,\varphi)}{\cos\frac{\theta}{2}}}\,q(t)\cdot\frac{\pi^2+t^2}{2}, \quad (6.19)$$

where the bar denotes averages over the solid angle.

The equation that has been obtained is complicated, but analysis shows that for arbitrary assumptions about the form of $w(\theta,\varphi)$ the error in the coefficient of viscosity will be smaller than 10 percent if we simply assume that the values we need satisfy $t^2 \ll \pi^2$. Then q is a constant, given by

$$q = \frac{64\pi^2\hbar^6}{3m^{*3}T^2}\left[\overline{\frac{w(\vartheta,\varphi)}{\cos\left(\frac{\theta}{2}\right)}(1-\cos\theta)^2\sin^2\varphi}\right]^{-1}. \quad (6.20)$$

The flux of momentum is given by (5.3). Inserting in it $n = n_0 + \delta n$ and Eq. (6.10), and using the fact that the energy is also a functional of the distribution function, we get:

$$\Pi_{ik} = \int p_i \frac{\partial\varepsilon(p)}{\partial p_k}\left[\delta n - \frac{\partial n_0}{\partial\varepsilon_0}\int f(\mathbf{p},\mathbf{p}')\,\delta n'\,d\tau'\right]d\tau = \int p_i \frac{\partial\varepsilon(p)}{\partial p_k}\frac{\partial n_0}{\partial\varepsilon_0}\left[\nu+\int f\nu\left(\frac{d\tau'}{d\varepsilon_0'}\right)_{\varepsilon_0'=\mu}\frac{d\Omega'}{4\pi}\right]d\tau \quad (6.21) = \int p_i \frac{\partial\varepsilon(p)}{\partial p_k}\frac{\partial n_0}{\partial\varepsilon_0}\,\psi\,d\tau.$$

Thus it is seen that the expression for the momentum flux involves the same function ψ that appears in the kinetic equation. Consequently, the function f does not enter the expression for the momentum flux, and it has the same form as for a Fermi gas of particles with the mass m^* and the scattering law described by the function $w(\theta,\varphi)$.*

Substituting Eqs. (6.17) and (6.20) in Eq. (6.21) and defining the coefficient of viscosity as the proportionality coefficient between Π_{ik} and $-(\partial u_i/\partial x_k + \partial u_k/\partial x_i - 2/3\ \partial u_l/\partial x_l\,\delta_{ik})$, we find

$$\eta = \frac{64}{45}T^{-2}\frac{\hbar^3 p_0^5}{m^{*4}}\left[\overline{\frac{w(\vartheta,\varphi)}{\cos\frac{\theta}{2}}(1-\cos\theta)^2\sin^2\psi}\right]^{-1}. \quad (6.22)$$

The viscosity is proportional to T^{-2}. This dependence was earlier predicted by Pomeranchuk[6] on the basis of qualitative considerations. As for the numerical values of the viscosity, they depend on the definite form of the averaged function $w(\theta,\varphi)$ and therefore cannot be found precisely. But Eq. (6.22) makes it possible to determine the order of magnitude of η. To do this we make use of the fact, which will be proved in Appendix 2, that the function f is of the order of magnitude of the scattering amplitude of the quasi-particles. Consequently $w \sim (2\pi/\hbar)f^2$, and the value of f can be determined from experimental data on the speed of sound (cf. Sec. 8). Using the numerical values of m^* and p_0 from Eq. (3.4), we get for He^3

$$\eta = \frac{\alpha}{T^2}\ (\alpha \sim 10^{-6} \div 10^{-5}\ \text{poise};\ T \text{ в } ^\circ\text{K}). \quad (6.23)$$

This order of magnitude corresponds to the experimental results found by K. N. Zinov'eva.[7] Because her measurements went down only to 0.35°K, it is indeed impossible to check the temperature dependence. Nevertheless the experimental results show an increase of the viscosity with decreasing temperature.

As already noted early in this section, apart from terms of order $(T/\mu)^2$ there is no part pro-

*In reference 5 the second term in the formula (6.9) was not included, and incorrect formulas were thus obtained for the viscosity and thermal conductivity coefficients. The correct values can be obtained by taking f = 0 in the expressions (26) and (30) of reference 5.

portional to div **u** on the left side of the kinetic equation. This means that the coefficient of second (dilatational) viscosity is smaller than η by a factor $(T/\mu)^2$ and plays no part in physical effects (for example, the absorption of sound).

7. THE THERMAL CONDUCTIVITY

We now go on to the calculation of the thermal conductivity.[5] Assuming a small temperature gradiant, we get in the left side of the kinetic equation:

$$-\frac{\partial n_0}{\partial \varepsilon_0}\left(\frac{\varepsilon_0-\mu}{T}-s\right)\frac{\partial \varepsilon_0}{\partial \mathbf{p}}\nabla T. \tag{7.1}$$

At the beginning the transformations of the collision integral are the same as those made in the preceding section [Eqs. (6.8) to (6.16)]. In the present case it is reasonable to try to find the function ψ in the form

$$\psi = q(t)\frac{\partial \varepsilon_0}{\partial \mathbf{p}}\nabla T. \tag{7.2}$$

Instead of Eq. (6.8) we now get a similar equation, but with $\cos\theta$ instead of $P_2(\cos\theta)$ in the collision integral and the factor $(\epsilon_0-\mu)/T - s$ in the left member. The presence of such a factor shows that the unknown function must contain parts both symmetric and antisymmetric in t, namely:

$$q(t) = q_s(t) + q_a(t). \tag{7.3}$$

Substitution in the collision integral gives the two equations

$$\frac{8\pi^4\hbar^6}{m^{*3}T^2}s = \frac{\overline{w(\theta,\varphi)}}{\cos\frac{\theta}{2}}\Big[q_s(t)\frac{\pi^2+t^2}{2} - \int_0^\infty dx\,\frac{x[q_s(x+t)+q_s(x-t)]}{e^x-1} - 2\int_0^\infty \frac{dx\,xq_s(x)}{e^x+1} - \int dx\,xq_s(x-t)\Big], \tag{7.4}$$

$$\frac{8\pi^4\hbar^6}{m^{*3}T^2} = \frac{\overline{w(\theta,\varphi)(1+2\cos\theta)}}{\cos\frac{\theta}{2}}\Big[\int_0^\infty dx\,\frac{x[q_a(x-t)-q_a(x+t)]}{e^x-1} \tag{7.5}$$

$$+2t\int_0^\infty \frac{dx\,q_a(x)}{e^x+1} + \int_0^t dx\cdot xq_a(x-t)\Big] + \frac{\overline{w(\theta,\varphi)}}{\cos\frac{\theta}{2}}q_a(t)\frac{\pi^2+t^2}{2}.$$

In the present case, however, the kinetic equation does not completely determine the solution of the problem. One must use along with it a supplementary condition, which expresses the conservation of the flux of matter. This is

$$\delta\int n\frac{\partial\varepsilon}{\partial\mathbf{p}}d\tau = \int \delta n\frac{\partial\varepsilon_0}{\partial\mathbf{p}}d\tau + \int n_0\frac{\partial}{\partial\mathbf{p}}\Big[\int f\,\delta n'\,d\tau'\Big]d\tau$$
$$= \int\frac{\partial\varepsilon_0}{\partial\mathbf{p}}\Big[\delta n - \frac{\partial n_0}{\partial\varepsilon_0}\int f\,\delta n'\,d\tau'\Big]d\tau = 0. \tag{7.6}$$

Here we have performed an integration by parts and have replaced ϵ by ϵ_0 in the integrals that already contain δn. Substituting the expressions (6.10), (6.11), and (7.2) in Eq. (7.6), we get

$$\int\frac{\partial n_0}{\partial\varepsilon_0}\Big(\frac{\partial\varepsilon_0}{\partial\mathbf{p}}\Big)^2 q(t)\,d\tau = 0. \tag{7.7}$$

To find $q_a(t)$ we have only to solve (7.5) in a way like that used in the treatment of the viscosity. The situation is different for the even part $q_s'(t)$. First, we must note that q_s = constant makes the right member of Eq. (7.4) vanish identically. Therefore the constant term $q_s(0)$ is to be determined not from this equation, but from the condition (7.7). Furthermore it is easy to show that the subsequent terms in $q_s(t)$, namely $a_2t^4 + a_4t^4 + \ldots$, make no contribution to the thermal conductivity. In fact, the occurrence of such a term as $a_m t^{2m}$, for example, must change the constant term in $q_s(t)$ by a quantity a_m^0 which is such that

$$\int\frac{\partial n_0}{\partial\varepsilon_0}\Big(\frac{\partial\varepsilon}{\partial\mathbf{p}}\Big)^2[a_m t^{2m}+a_m^0]\,d\tau = 0.$$

From this condition we get:

$$a_m^0 = -a_m(2m)!\,R_m;\quad R_m = \int_0^\infty \frac{z^{2m-1}}{e^z+1}\,dz.$$

To find the energy flux we have to calculate integrals of the form

$$\int\frac{\partial n_0}{\partial\varepsilon_0}F(\varepsilon_0)(a_m t^{2m}+a_m^0)\,d\tau$$
$$= \Big(F\frac{d\tau}{d\varepsilon_0}\Big)_{\varepsilon_0=\mu}[a_m(2m)!\,R_m + a_m^0] = 0.$$

Thus we get the result that to find the thermal conductivity coefficient we need only solve (7.5) and then choose the constant term so as to satisfy the condition (7.7). As in the case of the viscosity, it is enough for practical purposes to find the solution on the assumption $t^2 \ll \pi^2$. We then get:

$$q = \frac{24\pi^2\hbar^6}{m^{*3}T^2}\Big(t - \frac{\pi^2T_m^*}{p_0^2}\Big)\Big[\frac{\overline{w(\theta,\varphi)(1-\cos\theta)}}{\cos\frac{\theta}{2}}\Big]^{-1}. \tag{7.8}$$

The energy flux is given by Eq. (5.5). Substituting Eq. (6.1) in that formula, we have:

$$Q = \int\varepsilon_0(\mathbf{p})\frac{\partial\varepsilon_0(\mathbf{p})}{\partial\mathbf{p}}\delta n\,d\tau - \int\varepsilon_0(\mathbf{p})\frac{\partial\varepsilon_0(\mathbf{p})}{\partial\mathbf{p}}\cdot\frac{\partial n}{\partial\varepsilon_0}\int f\delta n'\,d\tau'\,d\tau. \tag{7.9}$$

Substituting Eqs. (6.10), (6.11), (7.2), and (7.8) we find the value of the thermal conductivity coefficient

$$\varkappa = \frac{8}{3}\frac{\pi^2\hbar^3 p_0^3}{m^{*4}T}\cdot\left[\frac{w(\theta,\varphi)(1-\cos\theta)}{\cos\frac{\theta}{2}}\right]^{-1}. \qquad (7.10)$$

As in the case of the viscosity, the function f does not appear in the coefficient κ.

The temperature dependence of the thermal conductivity coefficient is given by the T^{-1} law, which also agrees with the qualitative prediction of Pomeranchuk.

Estimating the numerical value of κ for He^3 in the same way as we did that of the viscosity, we get the relation

$$\varkappa = \frac{\beta}{T}\ (\beta \sim 10^2 \div 10^3 \text{ erg cm}^{-1}\text{ sec}^{-1}\text{ deg}^{-1};\ T \text{ in } {}^\circ\text{K}). \qquad (7.11)$$

In any case the expressions obtained for the viscosity and thermal conductivity coefficients cease to be valid for temperatures $T \sim \mu$. In addition to this, however, there is a limitation that shifts the region of applicability of the theory toward considerably lower temperatures. This is that the range of excitation energies with which we are concerned, of the order of T, must be much larger than the quantum uncertainty in the energy that arises owing to the collisions, i.e.,

$$\tau \gg \frac{\hbar}{T}, \qquad (7.12)$$

where τ is the time between collisions. We note that the condition (7.12) is required not only for the calculation of the kinetic coefficients as it was done above, but also for the validity of the entire theory of the Fermi fluid. For He^3 the quantity τ can be determined from the experimental values of the viscosity[7] [cf. Eq. (9.12)]. If we use the point at the very lowest temperature, i.e., $\eta = 4.8 \times 10^{-5}$ poise at $T = 0.35°K$, we get from Eq. (7.12)

$$T \ll 0{,}35^\circ\text{K}. \qquad (7.13)$$

It is clear, however, that this value is too high, since at $T = 0.35°K$ the viscosity is still not proportional to T^{-2}, but follows a weaker law of variation.

8. THE PROPAGATION OF SOUND

The propagation of sound in a Fermi fluid has a number of quite specific peculiarities. If we consider sound of a prescribed frequency, then at not too low temperatures its propagation takes place according to the laws of ordinary hydrodynamics. In this case the damping of the sound is proportional to τ. When the temperature is lowered the collision time will increase proportional to T^{-2}, and finally, at temperatures such that $\omega\tau \sim 1$, the sound ceases to be propagated at all. It turns out, however, that when the temperature is lowered still further the possibility of propagation of sound reappears. The velocity will then be different, and the sound is no longer simply a wave of compression and rarefaction. This phenomenon was also predicted by Landau,[8] and was called by him the "zeroth sound". Because only the relation between ω and τ is essential in fixing the nature of the sound, these two kinds of sound can be described as low-frequency sound ($\omega\tau \ll 1$) and high-frequency sound ($\omega\tau \gg 1$).

The speed of sound at low frequencies or not too low temperatures, for which the condition $\omega\tau \ll 1$ holds, is determined by the compressibility in the usual way. It was shown by Landau[1] that there is an important dependence of the speed on the function f.

It is convenient to express the compressibility in terms of the derivative $\partial\mu/\partial N$. Here N is the total number of particles. Using the fact that the chemical potential μ depends only on N/V, we find:

$$\frac{\partial\mu}{\partial N} = \frac{V\frac{\partial\mu}{\partial V}}{N} = -\frac{V^2}{N^2}\frac{\partial p}{\partial V}. \qquad (8.1)$$

As is well known, the speed of sound is given by the relation

$$c^2 = \frac{\partial p}{\partial\rho} = \frac{\partial p}{\partial\left(\frac{mN}{V}\right)} = \frac{1}{m}\left(N\frac{\partial\mu}{\partial N}\right). \qquad (8.2)$$

The derivative $\partial\mu/\partial N$ is calculated in the following way. Because $\mu = \epsilon(p_0)$, the change of μ occurs both on account of the change of p_0 and also because of the change of the form of the function $\epsilon(p)$:

$$\delta\mu = \int f\,\delta n'\,d\tau' + \frac{\partial\epsilon_0}{\partial p_0}\delta p_0. \qquad (8.3)$$

The changes δN and δp are connected by the relation

$$\delta N = 8\pi p_0^2\,\delta p_0\frac{V}{(2\pi\hbar)^3}. \qquad (8.4)$$

Since only changes δn close to the Fermi limit are important in the integral of Eq. (8.3), the integration over the absolute value of the momentum can be performed. This gives

$$\int f\,\delta n'\,d\tau' = \frac{\delta N}{4\pi V}\int f\,d\Omega. \qquad (8.5)$$

Substitution in Eq. (8.3) gives [cf. also Eqs. (2.1) and (2.2)]:

$$\frac{\partial\mu}{\partial N} = \frac{1}{4\pi V}\int f\,d\Omega + \frac{(2\pi\hbar)^3}{8\pi p_0 m^* V}. \qquad (8.6)$$

Equation (2.5) enables us to express the effective mass m^* appearing here in terms of the actual

mass m of the atoms. Substituting the resulting expression in Eq. (8.2), and using the fact that

$$N=\frac{8\pi p_0^3 V}{3(2\pi\hbar)^3},$$

we find:

$$c^2=\frac{p_0^2}{3m^2}+\frac{2}{3m}\left(\frac{p_0}{2\pi\hbar}\right)^3\int f(\chi)(1-\cos\chi)\,d\Omega. \quad (8.7)$$

The formulas (2.5) and (8.7) make it possible to get some information about the function $f(\chi)$. If we introduce the notation

$$F(\chi)=f(\chi)\left(\frac{d\tau}{d\varepsilon}\right)_{\varepsilon=\mu}=f(\chi)\frac{p_0 m^*}{\pi^2\hbar^3}=\sum_k F_k P_k(\cos\chi), \quad (8.8)$$

the first coefficients of the expansion in Legendre polynomials can be found from Eqs. (2.5) and (8.7). Indeed, it is not hard to get the relations

$$\frac{1}{m}=\frac{1}{m^*}(1+\overline{F\cos\chi}),\quad c^2=\left(\frac{p_0^2}{3m^2}\right)\frac{1+\overline{F}}{1+\overline{F\cos\chi}}. \quad (8.9)$$

Substituting the values of the parameters for He^3 from Eq. (3.4) and the value of the speed of sound,[9] $c_{T=0}=183$ m/sec, we find: d:

$$F_0=\overline{F}=4.65,\quad F_1=3\overline{F\cos\chi}=1.3. \quad (8.10)$$

In view of the fact that the later coefficients cannot be determined in such a simple way, hereafter we shall confine ourselves in specific cases to the first two harmonic contributions to the function F:

$$F=F_0+F_1\cos\chi. \quad (8.11)$$

We now go on to the study of the "zeroth sound", i.e., the case in which $\omega\tau\gg 1$. The kinetic equation has the form (5.4). We substitute in it the distribution function in the form

$$n=n_0+\delta n, \quad (8.12)$$

where the small deviation $\delta n\sim e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$. Here, of course, we have to take into account the fact that, according to Eq. (1.5), the energy of the excitations also undergoes a change. Inasmuch as the collision integral has the behavior $I(n)\sim\delta n/\tau$, and $\omega\tau\gg 1$, we can neglect I(n).

Thus we find that

$$(\mathbf{k}\cdot\mathbf{v}-\omega)\delta n-\mathbf{k}\cdot\mathbf{v}\frac{\partial n_0}{\partial\varepsilon}\int f\delta n'\,d\tau'=0.$$

From the form of this equation it follows that δn is proportional to $\partial n_0/\partial\epsilon$. Writing $\delta n=\partial n_0/\partial\epsilon\,\nu$, we get:

$$(\mathbf{k}\cdot\mathbf{v}-\omega)\nu+\mathbf{k}\cdot\mathbf{v}\int F\nu'\frac{d\Omega'}{4\pi}=0. \quad (8.13)$$

If we take $\mathbf{k}$ as the polar axis and introduce the symbols $u=\omega/k$ for the speed of propagation of the wave and $s=u/v$, then Eq. (8.13) takes the form:

$$(s-\cos\theta)\nu(\theta,\varphi)=\cos\theta\int F(\chi)\nu(\theta',\varphi')\frac{d\Omega}{4\pi}. \quad (8.14)$$

From this we can determine $\nu(\theta,\varphi)$ and the characteristic value s. In view of the fact that the change of the distribution function, δn, is proportional to $\partial n_0/\partial\epsilon$, it is clear that the essential point is the deformation of the Fermi surface. The shape of this deformation is determined by the function ν. Equation (8.14) depends essentially on the function F.

Let us consider first as an example the simplest case, namely $F=F_0=\text{const}$. With this we get from Eq. (8.14) (without the periodic exponential factor)

$$\nu=\frac{\text{const}\cdot\cos\theta}{s-\cos\theta}. \quad (8.15)$$

A displacement of the Fermi surface as a whole (this corresponds to ordinary sound) would be given by $\nu\sim\cos\theta$. Thus in the present case we have to do with an anisotropic deformation of the Fermi surface. As we shall see latter, s must be larger than unity. Thus the Fermi surface is stretched out in the direction of the motion.

Substituting Eq. (8.15) in Eq. (8.14) with $F=F_0$, we find the equation for s. Carrying out the elementary integration, we find:

$$\frac{s}{2}\ln\frac{s+1}{s-1}-1=\frac{1}{F_0}. \quad (8.16)$$

From this it can be seen that if s is real (which corresponds to undamped waves) it must be larger than 1, i.e.,

$$u>v. \quad (8.17)$$

From Eq. (8.14) it can be seen that this condition remains valid for any function F. Furthermore, since the left member of Eq. (8.16) is always positive, it is clear that the condition for the existence of zeroth sound is $F_0>0$.

If the function F_0 is large, s is also large. From Eq. (8.16) we get $s\to(F_0/3)^{1/2}$ for $F_0\to\infty$. On the other hand, for $F_0\to 0$ we have $s\to 1$, i.e., $u\to v$. This is the case of an almost free Fermi gas.

It is not hard to see that the conclusion that $s\to 1$ for $F\to 0$ does not depend on the form of F. In fact, it follows from Eq. (8.14) that for $F\to 0$, $s\to 1$, and ν is different from zero only for small θ. According to Eq. (8.7), in a weakly nonideal Fermi gas $c^2\approx p_0^2/3m$, i.e., $c\cong 3^{-1/2}u$. Thus the speed of the zeroth sound will exceed the ordinary speed of sound by a factor $3^{1/2}$.

Let us now consider the general case $F(\chi)$.[10] We represent this function as a sum of spherical har-

monics as in Eq. (8.8). We substitute Eq. (8.8) in Eq. (8.14) and use the addition theorem of Legendre polynomials

$$P_n(\chi) = \sum_{m=-n}^{n} P_n^m(\theta)\, P_n^m(\theta')\, e^{im(\varphi-\varphi')}\, \frac{(n-|m|)!}{(n+|m|)!},$$

where $P_n^m = P_n^{-m}$ are the associated Legendre polynomials. After the indicated substitution we get:

$$(\cos\theta - s)\,\nu - \cos\theta \sum \frac{(n-|m|)!}{(n+|m|)!} P_n^m(\theta) F_n e^{im\varphi} \times \int P_n^m(\theta')\,\nu(\theta', \varphi')\, e^{-im\varphi'} \frac{d\Omega'}{4\pi} = 0. \tag{8.18}$$

We introduce the notation

$$F_n \frac{(n-|m|)!}{(n+|m|)!} \int P_n^m(\theta')\,\nu(\theta'\varphi')\, e^{-im\varphi'} \frac{d\Omega'}{4\pi} = \Phi_{nm} \tag{8.19}$$

and solve Eq. (8.18) for ν:

$$\nu = \frac{\cos\theta}{\cos\theta - s} \sum \Phi_{nm} P_n^m(\theta)\, e^{im\varphi}. \tag{8.20}$$

Substituting this expression into Eq. (8.19) and carrying out the integration over φ we get:

$$F_n \frac{(n-|m|)!}{(n+|m|)!} \int \sum_k P_n^m(\theta') \frac{\cos\theta'}{\cos\theta' - s} P_k^m(\theta') \frac{d\Omega'}{4\pi} \Phi_{km} = \sum_k \Phi_{km}\delta_{kn}. \tag{8.21}$$

Thus we have obtained a system of homogeneous equations for the quantities Φ_{km}. This system separates into independent subsystems corresponding to different values of m. From Eq. (8.21) it follows that in the Fermi fluid at absolute zero vibrations of several different types can be propagated, characterized by different dependences of the amplitude on the angles θ, φ. The value m = 0 corresponds to vibrations in which ν is isotropic in the plane perpendicular to **k**. For $m \neq 0$ the vibrations are polarized in a definite way in this plane. The number of types of vibrations is determined by the number of possible values of m ($|m| \leq n$). The speeds of propagation of the vibrations are determined by the requirement that the determinants of the corresponding subsystems of equations be equal to zero:

$$\| \delta_{kn} + F_n \Omega_{kn}^m(s) \| = 0 \quad (N \geq n,\ k \geq |m|)$$
$$\Omega_{kn}^m(s) = \frac{(n-|m|)!}{(n+|m|)!} \int P_k^m(\theta') \frac{\cos(\theta')}{\cos\theta' - s} P_n^m(\theta') \frac{d\Omega'}{4\pi}. \tag{8.22}$$

Because $P_n^{-m} = P_n^m$, the coefficients Ω_{kn}^m do not depend on the sign of m, so that vibrations that differ only by the sign of m are propagated with the same speed.

From Eq. (8.22) it can be seen that the equations for the speeds are transcendental equations. In the general case they do not always have real roots. Cases are also possible, however, in which there are several real roots. Then there are several types of vibrations with the same polarization in the plane perpendicular to **k**.

As an example let us consider the case in which the function $F(\chi)$ contains only the zeroth and first harmonics [the function (8.11)]. Here the coefficients Ω_{kn}^m are given by

$$\Omega_{00}^0 = \frac{1}{2}\int_{-1}^{1} \frac{x\,dx}{x-s} = 1 - \frac{s}{2}\ln\frac{s+1}{s-1} = -w,$$

$$\Omega_{10}^0 = \Omega_{01}^0 = \frac{1}{2}\int_{-1}^{1} \frac{x^2 dx}{x-s} = -sw,$$

$$\Omega_{11}^0 = \frac{1}{2}\int_{-1}^{1} \frac{x^3\,dx}{x-s} = \frac{1}{3} - s^2 w,$$

$$\Omega_{11}^1 = \frac{1}{4}\int_{-1}^{1} \frac{(1-x^2)}{x-s}\,dx = \frac{1}{2}\left[(s^2-1)w - \frac{1}{3}\right].$$

On substituting in the determinant (8.22) we get for the speed of propagation of vibrations of type m = 0 the equation

$$w = \frac{1 + F_1/3}{F_0 + F_0 F_1/3 + F_1 s^2}. \tag{8.23}$$

For the case m = 1 we get the equation

$$w = \frac{F_1 - 6}{3F_1(s^2-1)}. \tag{8.24}$$

This equation has one real root for $F_1 > 6$.

Let us now look at the application of our formulas to the case of liquid He^3. According to Eq. (8.10) there is one speed for the vibrations of type m = 0 [the root of Eq. (8.23)]

$$s = \frac{u}{v} = 1.72, \quad u = 192 \text{ m/sec}.$$

The vibrations of type m = 1 (and also all with m > 1) are absent. It is of course possible that this conclusion is a consequence of the crudeness of the approximation we have taken for $F(\chi)$, but we see no reasons for thinking so.

In addition to the vibrations that have been considered, waves of an entirely different type can be propagated in a Fermi fluid at absolute zero; these can be called spin waves.[8] In the treatment of these vibrations one must take into account the spin part of the function f, which is given by Eq. (4.3). We introduce a new symbol

$$\Phi(\chi) = f(\chi)\left(\frac{dn}{d\varepsilon}\right)_{\varepsilon=\varepsilon_F} = F(\chi) + Z(\chi)\,\sigma\cdot\sigma'. \tag{8.25}$$

The equation for zeroth sound now takes the form

$$(s-\cos\theta)\,\nu=\cos\theta\cdot\frac{1}{2}\,\mathrm{Sp}_{\sigma'}\frac{1}{4\pi}\int\Phi\nu'\,d\Omega'. \qquad (8.26)$$

If we assume that ν does not depend on the spin, this goes over into Eq. (8.14). It is also possible, however, to assume a different form for the function ν, namely to take

$$\nu=\boldsymbol{\nu}\cdot\boldsymbol{\sigma} \qquad (8.27)$$

For this function we get the equation

$$(s-\cos\theta)\,\boldsymbol{\nu}=\cos\theta\int z\boldsymbol{\nu}'\frac{d\Omega'}{16\pi}. \qquad (8.28)$$

Thus we have for the components of the vector $\boldsymbol{\nu}$ an equation that differs from Eq. (8.14) only by the replacement of F by Z/4. Therefore all the further developments must be the same as those carried out above.

According to the results of Sec. 4, for the case of He^3 the quantity Z_0 is approximately -3.4, that is, a negative quantity. Although it is of course impossible to draw any rigorous conclusions from this, still it is most probable that no spin waves can be propagated in He^3.

The possibility of the propagation of sound waves at T = 0 means that the Fermi fluid has a Bose branch of the energy spectrum, in the form of phonons with energy ϵ = up. But the contributions from this branch to the thermodynamic quantities involve higher powers of T (heat capacity $\sim T^3$), which were not included in the approximate theory that has been considered.

9. THE DISPERSION AND ABSORPTION OF SOUND

In the preceding section we considered the two limiting cases $\omega\tau\ll1$ and $\omega\tau\gg1$. Let us now examine how the transition occurs from ordinary sound to zeroth sound, i.e., let us consider the dispersion of sound.[10] Because the region of the dispersion, $\omega\tau\sim1$, is at the same time the region of strong absorption, our treatment will also enable us to deal with the problem of the absorption of sound.

In the present case we have to find the solution of the kinetic equation, including the collision integral. Because the interaction law of the excitations is still unknown, however, we shall simplify the form of the collision integral.

For this purpose we could introduce a certain effective time τ and replace the collision integral I in Eq. (5.1) by the expression $-\delta n/\tau$. With this replacement, however, the kinetic equation will not yield any conservation laws for the number of excitations, the momentum, and the energy, and this makes the transition to hydrodynamics impossible. Since both the second (dilatational) viscosity and the effects of thermal conductivity (see below) are negligibly small in our case, the collision integral contains no terms in the zeroth and first spherical harmonics. We therefore replace the collision integral by the following expression:*

$$I(n)\to-\tau^{-1}(\delta n-\overline{\delta n}-\overline{3\delta n\cos\theta}\cos\theta). \qquad (9.1)$$

It is easy to see that the integration over $d\tau_p$ makes this expression go to zero. It also goes to zero after multiplication by ϵ or by p cos θ and integration over $d\tau_p$.† Thus the equations for the conservation of the number of particles, the momentum, and the energy are automatically satisfied. According to Eqs. (8.14) and (9.1) the basic kinetic equation now takes the following form:

$$(kv\cos\theta-\omega)\,\nu+kv\cos\theta\int F\nu'\frac{d\Omega'}{4\pi}=-\frac{1}{i\tau}(\nu-\bar\nu-\overline{3\nu\cos\theta}\cdot\cos\theta). \qquad (9.2)$$

Having in mind the application of the results to He^3, in order not to complicate the problem we shall write the function F in the two-term form of Eq. (8.11).

We introduce the notation

$$\bar\nu=\nu_0;\quad\overline{3\nu\cos\theta}=\nu_1;\quad\sigma=-i\tau kv;\quad\xi=\frac{i\omega\tau-1}{i\tau kv}, \qquad (9.3)$$

after which we get without difficulty from Eq. (9.2):

$$(\cos\theta-\xi)\,\nu+\cos\theta\left(F_0\nu_0+\frac{1}{3}F_1\nu_1\cos\theta\right)=\frac{1}{\sigma}(\nu_0+\nu_1\cos\theta). \qquad (9.4)$$

We now solve the equation for ν and calculate $\bar\nu=\nu_0$ and $\overline{\nu\cos\theta}=\nu_1/3$. We thus find two equations for the two quantities ν_0 and ν_1:

$$\nu_0=F_0\nu_0w+\frac{1}{3}wF_1\nu_1-\frac{1}{\sigma}\frac{w+1}{\xi}\nu_0-\frac{\nu_1}{\sigma}w, \qquad (9.5)$$

$$\frac{1}{3}\nu_1=F_0\nu_0\xi w-\frac{1}{3}\left(\frac{1}{3}-\xi^2w\right)F_1\nu_1-\frac{w}{\sigma}\nu_0-\frac{\xi w}{\sigma}\nu_1, \qquad (9.6)$$

where $w=\frac{1}{2}\ln\frac{\xi+1}{\xi-1}-1$.

From the condition for compatibility of (9.5) and (9.6) we get an equation for the complex speed of sound

$$\left(1+\frac{1}{\xi\sigma}\right)\left(1+\frac{F_1}{3}\right)-w\left\{\left(1+\frac{F_1}{3}\right)\left(F_0-\frac{1}{\xi\sigma}\right)+\xi^2\left(F_1-\frac{3}{\xi\sigma}\right)\cdot\left(1+\frac{1}{\xi\sigma}\right)\right\}=0. \qquad (9.7)$$

This is the desired equation for the frequency

*Here and in what follows the bar denotes averaging over the angles.

†In this integration only the range of values of momentum close to p_0 is of importance, since the function δn contains a δ-function singularity for $\epsilon=\mu$.

dependence of the speed of sound or, in other words, for the dispersion of sound in the Fermi fluid. We first consider two extreme cases.

a) Low frequencies, $\omega\tau \ll 1$. Here we have $\sigma \to 0$, $\xi\sigma \to -1$, and $\xi \to \infty$. Expansion of w in powers of $1/\xi$ gives

$$w = \frac{1}{3\xi^2} + \frac{1}{5\xi^4}$$

and after simple manipulations Eq. (9.7) takes the form

$$\left(1+\frac{1}{\xi\sigma}\right)^2 = \frac{1}{3\xi^2}\cdot\left(1+\frac{F_1}{3}\right)\cdot\left(F_0-\frac{1}{\xi\sigma}\right) + \frac{1}{5\xi^2}\left(1+\frac{1}{\xi\sigma}\right)\cdot\left(F_1-\frac{3}{\xi\sigma}\right). \tag{9.8}$$

Recalling the relations (9.3), we have:

$$\left(\frac{1+\xi\sigma}{\sigma}\right)^2 = \left(\frac{\omega}{kv}\right)^2.$$

From Eqs. (9.8) and (9.3) we find in first order in $i\omega\tau$

$$\left(\frac{\omega}{kv}\right)^2 = \frac{1}{3}(1+F_0)\left(1+\frac{F_1}{3}\right) - \frac{4}{15}i\omega\tau\left(1+\frac{F_1}{3}\right). \tag{9.9}$$

The first term corresponds to the speed of ordinary sound in the Fermi fluid. The absorption of sound in the region $\omega\tau \ll 1$ is found by the elementary rule from Eq. (9.9) as the imaginary part of the wave vector:

$$\gamma = \operatorname{Im} k = \frac{2\omega^2\tau v^2}{15c^3}\left(1+\frac{F_1}{3}\right). \tag{9.10}$$

In view of the fact that the present case corresponds to ordinary sound, we can use instead of Eq. (9.10) the more precise ordinary formula

$$\gamma = \frac{\omega^2}{2\rho c^3}\left\{\left(\frac{4}{3}\eta+\zeta\right)+\frac{\varkappa}{c_p}\left(\frac{c_p}{c_v}-1\right)\right\}.$$

Noting that $\xi \ll \eta$ and also that κ/c_p, which is a quantity of the order of η, is multiplied by the small factor $(c_p/c_v) - 1$ [it is not hard to verify that it is of the order $(T/\mu)^2$], we find:

$$\gamma = \frac{2\omega^2}{3\rho c^3}\eta. \tag{9.11}$$

Comparing this expression with Eq. (9.10), we can determine the time τ from experimental data:

$$\eta = \frac{1}{5}\rho\tau v^2\left(1+\frac{1}{3}F_1\right). \tag{9.12}$$

From the data of K. N. Zinov'eva on the viscosity[7] we get

$$\tau \cong 2.3\cdot 10^{-12}\cdot T^{-2} \text{ sec } (T \text{ in } ^\circ\text{K}) \tag{9.13}$$

The corresponding value for the absorption coefficient is

$$\gamma \sim 1\cdot 10^{-17}\left(\frac{\omega}{T}\right)^2 \text{cm}^{-1}. \tag{9.14}$$

b) We now consider the second extreme case, that of high frequencies and low temperatures, $\omega\tau \gg 1$. In this case

$$\sigma \to \infty;\ \xi\sigma \to \infty;\ \xi = s + i\xi',\ |\xi'| \ll s.$$

Equation (9.7) takes the form

$$\left(1+\frac{F_1}{3}\right) - w(s)\left\{\left(1+\frac{F_1}{3}\right)F_0 + s^2F_1\right\} = 0. \tag{9.15}$$

This equation agrees exactly with the equation (8.23) that determines the speed of the zeroth sound.

As for the absorption of the zeroth sound, to calculate it we must find the imaginary part ξ' of the speed of sound. From Eq. (9.9) we get the equation

$$\xi'\left\{\frac{1}{w(s)}\left(1+\frac{F_1}{3}\right)\left(\frac{s}{s^2-1}-\frac{w(s)+1}{s}\right) - 2sw(s)F_1\right\} - \frac{1}{\omega\tau}\left\{\left(1+\frac{F_1}{3}\right)(1+w(s)) + w(s)s^2(3-F_1)\right\} = 0. \tag{9.16}$$

By means of Eq. (9.3) we find the absorption coefficient

$$\gamma = \operatorname{Im} k = \frac{1}{s\tau} - \frac{\omega\xi'}{s^2}. \tag{9.17}$$

If we substitute here the values of the parameters for He^3 ($s = 1.84$, $v = 1.13 \times 10^4$ cm/sec), we get:

$$\xi' = \frac{1.5}{\omega\tau},\quad \gamma \sim 3\cdot 10^6\cdot T^2 \text{cm}^{-1}. \tag{9.18}$$

Thus the absorption of the zeroth sound does not depend on the frequency and increases with increasing temperature as $1/\tau$, i.e., proportional to T^2.

In this calculation the speed of sound has been treated classically. The validity of this treatment depends on the inequality $\hbar\omega \ll T$. For the case $\hbar\omega \gtrsim T$ it is necessary to treat the problem by quantum theory. The decrease of the number of sound quanta per unit time, due to the collisions of the Fermi particles, is given by

$$\int w\{n_1n_2(1-n_1')(1-n_2') - n_1'n_2'(1-n_1)(1-n_2)\} \times \delta(\mathbf{p}'+\mathbf{p}_2'-\mathbf{p}_1-\mathbf{p}_2-\hbar\mathbf{k}) \times \delta(\varepsilon_1'+\varepsilon_2'-\varepsilon_1-\varepsilon_2-\hbar\omega)\, d\tau_1\, d\tau_2\, d\tau_1'\, d\mathbf{p}_2'. \tag{9.19}$$

The function w, which depends on the four momenta, is unknown. It is possible, however, to carry out the calculation in such a way that the absorption coefficient is expressed in terms of its classical value (9.17); this has been done in a paper by Landau.[8]

Making use of the fact that in the region in which the Fermi distribution falls off (which is the impor-

tant one for the problem) the only rapidly varying function is the occupation number $n(\epsilon)$, we can assume that the angular integrals give a certain constant, which does not change as we go from the classical region $\hbar\omega \ll T$ to the quantum region $\hbar\omega \gg T$. There then remains only the integral over the energies

$$I = \int \{n_1 n_2 (1-n_1')(1-n_2') - n_1' n_2' (1-n_1)(1-n_2) \times \delta(\varepsilon_1' + \varepsilon_2' - \varepsilon_1 - \varepsilon_2 - \hbar\omega)\, d\varepsilon_1\, d\varepsilon_2\, d\varepsilon_1'\, d\varepsilon_2'. \tag{9.20}$$

Substituting the Fermi function for n and extending the integration from $-\infty$ to ∞ (this is permissible because of the rapid convergence of the integrals), after a rather lengthy calculation we get

$$I = T^3 \frac{2\pi^2}{3} \hbar\omega \left[1 + \left(\frac{\hbar\omega}{2\pi T}\right)^2\right]. \tag{9.21}$$

The absorption coefficient is proportional to the integral I. From the classical limiting case we can determine the proportionality factor. Since the classical value of γ must be obtained for $\hbar\omega \ll T$, we have in the general case

$$\gamma = \gamma_{cl} \left[1 + \left(\frac{\hbar\omega}{2\pi T}\right)^2\right], \tag{9.22}$$

where γ_{cl} is given by Eq. (9.17). In view of the fact that γ_{cl} is proportional to T^2, in the quantum limiting case $\hbar\omega \gg T$ the absorption coefficient γ must be proportional to ω^2.

10. THE SCATTERING OF LIGHT (FLUCTUATIONS OF THE DISTRIBUTION FUNCTION)

According to the foregoing, a particular condition for the possibility of the propagation of "zeroth sound" is the inequality

$$\omega_{sound}\tau \gg 1, \tag{10.1}$$

where τ is the time between collisions. For He^3 this is of the order of $10^{-12}T^{-2}$ sec. If we take even a temperature $\sim 0.01°K$, a frequency higher than 10^8 cps is required for the direct observation of zeroth sound, and this makes the performance of such an experiment very difficult.

Besides the direct experiment, an indirect method can be proposed, which consists of the observation of the Rayleigh scattering of light in liquid He^3.* As is well known, in Rayleigh scattering there appear, in addition to the main line, satellites that differ from it in frequency by

$$\Delta\omega = \pm 2\frac{u}{c}\omega \sin\frac{\theta}{2},$$

where u is the speed of sound and θ is the scattering angle. The speed of the zeroth sound in He^3 is of the order of 2×10^4 cm/sec, that is, $\Delta\omega \sim 10^{-6}\omega$. Thus the observation of the frequency distribution of the scattered light provides in principle a possibility for measuring the speed of the zeroth sound. The condition (10.1) can be satisfied owing to the high frequency of visible light.

In addition to this side of the matter, the scattering of light in a Fermi fluid at sufficiently low temperatures has a number of specific features which give interest to the theoretical study of this phenomenon, in particular the frequency distribution of the intensity.*

As is well known, the frequency and angular distributions for the Rayleigh scattering of unpolarized light are given by the formula[12]†

$$dh = \frac{\omega^4}{6\pi c^4} \cdot \frac{1}{2\pi V} \overline{\left|\int \delta D_{\Delta\omega}(\mathbf{r})\, e^{-i\mathbf{q}\cdot\mathbf{r}} dV\right|^2} \frac{3}{4}(1+\cos^2\theta)\frac{d\Omega}{4\pi} d\Delta\omega, \tag{10.2}$$

where ω is the frequency of the incident light, θ is the scattering angle, $\mathbf{q}$ is the change of the wave vector of the light, of magnitude $(2\omega/c)\sin(\theta/2)$, and $\delta D_{\Delta\omega}$ is the Fourier component of the fluctuation $\delta D(t)$ of the dielectric constant:

$$\delta D_{\Delta\omega} = \frac{1}{\sqrt{t_0}} \int_0^{t_0} \delta D(t)\, e^{i\Delta\omega t} d\tau, \tag{10.3}$$

where t_0 is a certain large value, which will be let go to infinity in the final formula.

The bar in Eq. (10.2) means averaging over the fluctuations. Hereafter we shall for simplicity take the volume of the system to be unity.

Because of the very small polarizability of helium atoms we can assume that the change of the dielectric constant occurs as a result of density fluctuations, i.e., $\delta D = (\partial D/\partial N)\delta N$, where N is the number of particles in unit volume. But according to the general theory of the Fermi fluid the number of excitations is equal to the number of atoms in the fluid. Thus we can write

$$\int \delta D_{\Delta\omega}(\mathbf{r})\, e^{-i\mathbf{q}\cdot\mathbf{r}} dV = \frac{\partial D}{\partial N} \int \delta n_{\mathbf{q},\Delta\omega}(\mathbf{p})\, d\tau_p, \tag{10.4}$$

where $\delta n_{\mathbf{q},\Delta\omega}(\mathbf{p})$ is the Fourier component with respect to r and t [the latter in the sense of Eq. (10.3)] of the fluctuation of the distribution function of the excitations.

Before going on to further calculations, let us

*The idea of using the Rayleigh scattering was first proposed by S. P. Kapitza.

*We note that in the case of high temperatures, for which $\omega_{sound} \ll 1$, the scattering of light will be described by the usual formulas (cf. reference 12).

†Here dh is the so-called differential extinction coefficient. The integral of dh with respect to $d\Omega$ and $d\Delta\omega$ gives the total extinction coefficient h, which is the damping decrement of the photon flux density in the medium.

note one important fact. In Eq. (10.2) the averaging is taken over all possible fluctuations. In the range of temperatures and frequencies in which $\hbar\Delta\omega \gtrsim T$ we must take account of quantum effects in carrying out this averaging. For this, however, (cf. reference 12) it is enough to know the result for the purely classical case, that is, for $T \gg \hbar\Delta\omega$, and then introduce a certain correction factor. For scattering with increase of the frequency by the amount $\Delta\omega$ (anti-Stokes scattering) we must introduce the factor $(h\Delta\omega/T)N(\Delta\omega)$, and for scattering with decrease of the frequency by $\Delta\omega$ (Stokes scattering) the factor is $(h\Delta\omega/T)[N(\Delta\omega)+1]$, where $N(\Delta\omega)$ is the Bose distribution function. If we use negative values of $\Delta\omega$ for the description of the Stokes scattering, then owing to the relation $N(-\Delta\omega)+1 = -N(\Delta\omega)$ it turns out that the correction factor for both cases has the form

$$\frac{\hbar\Delta\omega}{T}\left[e^{\frac{\hbar\Delta\omega}{T}}-1\right]^{-1}. \tag{10.5}$$

Thus we shall take $T \gg \hbar\omega$. To find the fluctuations of the distribution function we shall use the method proposed by Rytov[13] and by Landau and Lifshitz[14] for calculating fluctuations in electrodynamics and hydrodynamics.* By this method we find the fluctuation of the "random force" appearing in the kinetic equation; then by solving this equation we can also get the fluctuation of the distribution function.

For the case of the Fermi fluid we shall start with the kinetic equation, which we write in the form

$$\frac{\partial\delta n}{\partial t}+\frac{\partial\delta n}{\partial\mathbf{r}}\cdot\frac{\partial\varepsilon}{\partial\mathbf{p}}-\frac{\partial n_0}{\partial\mathbf{p}}\int f(\mathbf{p},\mathbf{p}')\frac{\partial\delta n(\mathbf{p}')}{\partial\mathbf{r}}\,d\tau_{p'} = I(\delta n)+y(\mathbf{p},\mathbf{r},t). \tag{10.6}$$

Except for the "random force" $y(\mathbf{r}, \mathbf{p}, t)$ this equation is an approximation to the kinetic equation (5.1) that is linear in δn.

In what follows we shall be concerned only with the case of frequencies and temperatures satisfying the relation (10.1), i.e., with the case in which the collisions can be neglected. The detailed form of the collision integral is then of no importance, since it plays the part of an auxiliary quantity in the calculations and can be set equal to zero in the final result. In view of this we set

$$I(\delta n) = -\frac{\delta n}{\tau}, \tag{10.7}$$

where τ is a large quantity. We have also to find the rate of change of the entropy. Recalling that the number of particles and the total energy are prescribed, and using the relation

$$\delta\varepsilon(\mathbf{p}) = \int f(\mathbf{p},\mathbf{p}')\,\delta n(\mathbf{p}')\,d\tau_{p'}, \tag{10.8}$$

we find:

$$S = -\left\{\int\frac{\delta n\,[I(\delta n)+y]}{n_0(1-n_0)}\,d\tau_p\,dV + \frac{1}{T}\int f(\mathbf{p},\mathbf{p}')\,\delta(\mathbf{r}-\mathbf{r}')\,\delta n I(\delta n')\,d\tau_p\,dV\,d\tau_{p'}\,dV'\right\}. \tag{10.9}$$

Noting that $n_0(1-n_0) \approx T\delta(\epsilon-\mu)$, where μ is the chemical potential, we readily see that $\delta n(\mathbf{p})$ must have the form

$$\delta n(\mathbf{p}) = \nu(\theta,\varphi)\cdot\delta(\varepsilon-\mu), \tag{10.10}$$

where θ and φ are the polar angles of the vector $\mathbf{p}$. This formula means that the fluctuations of the distribution function occur only in the neighborhood of the Fermi surface.

It is natural to take a similar form for $\mathbf{y}$, i.e.,

$$y(\mathbf{p}) = y^{\varepsilon}(\theta,\varphi)\,\delta(\varepsilon-\mu). \tag{10.11}$$

We now introduce the notation

$$F(\chi) = \left[f(\mathbf{p},\mathbf{p}')\frac{d\tau}{d\varepsilon}\right]_{\varepsilon=\varepsilon'=\mu},$$

where χ is the angle between $\mathbf{p}$ and $\mathbf{p}'$, and expand δn, y, and F in spherical harmonics

$$\left.\begin{aligned}\nu(\theta,\varphi) &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n}A_n^m P_n^m(\cos\theta)\,e^{im\varphi},\\ y^{\varepsilon}(\theta,\varphi) &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n}y_n^m P_n^m(\cos\theta)\,e^{im\varphi},\\ F(\chi) &= \sum_{n=0}^{\infty}F_n P_n(\cos\chi).\end{aligned}\right\} \tag{10.12}$$

Since ν and y^{ϵ} are real quantities, we have

$$A_n^m = (A_n^{-m})^*,\quad y_n^m = (y_n^{-m})^*.$$

Using the expression (10.7) for the collision integral, we get the following formula for the rate of change of the entropy:

$$\dot S = \int dV\left(\frac{d\tau_p}{d\varepsilon}\right)_{\varepsilon=\mu}\frac{1}{T}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left(\frac{F_n}{2n+1}+1\right)\times\frac{1}{2n+1}\left(\frac{(n+|m|)!}{(n-|m|)!}\right)\left(\frac{A_n^m}{\tau}-y_n^m\right)A_n^{-m}. \tag{10.13}$$

Let us now introduce the notation

$$\dot x_n^m = -\frac{A_n^m}{\tau}+y_n^m. \tag{10.14}$$

Then in order for the expression (10.13) to take the form (cf. reference 14)

$$\dot S = -\sum_i X_i\cdot\dot x_j$$

*The writers express their gratitude to L. P. Gor'kov, I. E. Dzyaloshinskii, and L. P. Pitaevskii, who called attention to the possibility of applying this method to the kinetic equation.

we must take as the generalized forces X_i the expressions

$$X_n^m = \frac{1}{T}\left(\frac{d\tau_p}{d\varepsilon}\right)_{\varepsilon=\mu}\left(\frac{F_n}{2n+1}+1\right)\frac{1}{2n+1}\cdot\frac{(n+|m|)!}{(n-|m|)!}A_n^{-m}. \tag{10.15}$$

In the expression (10.14) the quantity y_n^m plays the role of the "random force." Writing this relation in the form

$$x_n^m = -\sum_{n'm'}\gamma_{n,n'}^{m,m'}X_{n'}^{m'} + y_n^m,$$

where the coefficients γ can be determined without difficulty from Eq. (10.15), we have according to the general theory of fluctuations:

$$\overline{y_n^m(\mathbf{r},t)\,y_{n'}^{m'}(\mathbf{r}',t')} = (\gamma_{n,n'}^{m,m'}+\gamma_{n',n}^{m',m})\,\delta(t-t')\,\delta(\mathbf{r}-\mathbf{r}')$$

$$= \frac{2}{\tau}\,\delta_{n,n'}\delta_{m,-m'}\,\delta(\mathbf{r}-\mathbf{r}')\,\delta(t-t') \tag{10.16}$$

$$\times\left[T\left(\frac{d\tau_p}{d\varepsilon}\right)_{\varepsilon=\mu}\left(\frac{F_n}{2n+1}+1\right)\frac{1}{2n+1}\frac{(n+|m|)!}{(n-|m|)!}\right]^{-1}.$$

Finally, using Eq. (10.12) and the relation

$$\sum_n (2n+1)\,P_n(\cos\chi) = 2\delta(\cos\chi-1),$$

we get after some transformations the general relation

$$\overline{y(\mathbf{p},\mathbf{r},t)\,y(\mathbf{p}',\mathbf{r}',t')} = \frac{2T}{\tau}\,\delta(\mathbf{r}-\mathbf{r}')\,\delta(t-t') \tag{10.17}$$

$$\times\left\{\frac{(2\pi\hbar)^3}{2}\,\delta(\mathbf{p}-\mathbf{p}')\,\delta(\varepsilon-\mu)+\delta(\varepsilon-\mu)\,\delta(\varepsilon'-\mu)\left(\frac{d\varepsilon}{d\tau_p}\right)_{\varepsilon=\mu}\right.$$

$$\left.\times\sum_{n=0}^{\infty}\frac{F_nP_n(\cos\chi)}{1+F_n/2n+1}\right.\cdot$$

By means of this formula and the kinetic equation (10.16) we can calculate the fluctuations of the distribution function, with which we are concerned here. Since in the general case of an arbitrary function f this is a rather complicated task, we confine ourselves to the case f = const.

Making use of the fact that the fluctuations occur only at the Fermi surface, and referring to Eq. (10.6), we express $\nu_{\mathbf{q},\Delta\omega}(\theta,\varphi)$ in terms of the corresponding Fourier component of $y^{\epsilon}(\theta,\varphi)$. This gives

$$\int \nu_{\mathbf{q},\Delta\omega}(\theta,\varphi)\,\frac{d\Omega}{4\pi}$$

$$= \int\frac{y^{\varepsilon}_{\mathbf{q},\Delta\omega}(\theta,\varphi)\,\frac{d\Omega}{4\pi}}{-i\omega+\frac{1}{\tau}+i\mathbf{q}\mathbf{v}}\Bigg/\left(1+F_0\int\frac{i(\mathbf{qv})\,\frac{d\Omega}{4\pi}}{-i\omega+\frac{1}{\tau}+i\mathbf{q}\mathbf{v}}\right), \tag{10.18}$$

where $\mathbf{v} = (\partial\epsilon/\partial\mathbf{p})_{\epsilon=\mu}$. Averaging the square of the magnitude of this expression by means of Eq. (10.17)[here we have to remember that the Fourier components with respect to t are taken according to Eq. (10.3)], we find:

$$\overline{\left|\int\nu_{\mathbf{q},\Delta\omega}(\theta,\varphi)\frac{d\Omega}{4\pi}\right|^2} = 2\left(\frac{d\varepsilon}{d\tau_p}\right)_{\varepsilon=\mu}\frac{T}{\tau}\left[\frac{1}{2}\int_{-1}^{1}\frac{dx}{\left|qvx-\omega+\frac{i}{\tau}\right|^2}\right. \tag{10.19}$$

$$\left.-\frac{F_0}{1+F_0}\left|\frac{1}{2}\int_{-1}^{1}\frac{dx}{qvx-\omega+\frac{i}{\tau}}\right|^2\right]\Bigg/\left|1+\frac{F_0}{2}\int_{-1}^{+1}\frac{qvx\,dx}{qvx-\omega+\frac{i}{\tau}}\right|^2.$$

We are interested in the limit of the value of this expression for $\tau\to\infty$. In the case in which $qv > |\Delta\omega|$, the denominator does not have any poles, and the important quantity is the residue in the integral in the numerator. We get as the result:

$$\frac{1}{2\pi}\overline{|\nu_{\mathbf{q},-\omega}(\mathbf{p})\,d\tau_p|^2} = T\left(\frac{d\tau_p}{d\varepsilon}\right)\frac{1}{qv}\left\{\left[1+F_0\left(1\right.\right.\right.$$

$$\left.\left.\left.-\frac{\Delta\omega}{2qv}\ln\frac{qv+\Delta\omega}{qv-\Delta\omega}\right)\right]^2+\left(\frac{F_0\Delta\omega\pi}{2qv}\right)^2\right\}^{-1}. \tag{10.20}$$

In the opposite case, i.e., for $qv \ll \Delta\omega$, the pole in the denominator of the expression (10.19) is important. It is not hard to see that such a pole occurs, if $F_0 > 0$ and is given by $\Delta\omega = \pm sqv$, where s satisfies the equation

$$1+F_0\left[1-\frac{1}{2}\ln\left(\frac{s+1}{s-1}\right)\right]=0. \tag{10.21}$$

This equation is identical with Eq. (8.16), which determines the speed of the zeroth sound. Using the relation

$$\frac{1}{\pi}\lim_{\tau\to\infty}\frac{\frac{1}{\tau}}{(\omega-\omega_0)^2+\tau^{-2}} = \delta(\omega-\omega_0),$$

we get without difficulty:

$$\frac{1}{2\pi}\overline{\left|\int\nu_{\mathbf{q},\Delta\omega}(\mathbf{p})\,d\tau_p\right|^2} = T\left(\frac{d\tau_p}{d\varepsilon}\right)_{\varepsilon=\mu}\frac{2(s^2-1)}{F_0(1+F_0-s^2)} \tag{10.22}$$

$$\times[\delta(\Delta\omega-sqv)+\delta(\Delta\omega+sqv)].$$

Thus the distribution of the scattered light in direction and frequency has the form [we have also introduced the quantum factor (10.5)]:

$$dh = \frac{\omega^4}{4\pi c^4}\left(\frac{\partial D}{\partial N}\right)^2\left(\frac{d\tau_p}{d\varepsilon}\right)_{\varepsilon=\mu}(1+\cos^2\theta)\cdot\frac{\hbar\Delta\omega}{e^{\hbar\Delta\omega/T}-1}$$

$$\times\left[\frac{\theta(qv-|\Delta\omega|)}{2qv}\right]\Bigg/\left\{\left[1+F_0\left(1-\frac{\Delta\omega}{2qv}\ln\frac{qv+\Delta\omega}{qv-\Delta\omega}\right)\right]^2\right. \tag{10.23}$$

$$\left.+\left(\frac{F_0\Delta\omega\pi}{2qv}\right)^2\right\}+\frac{s^2-1}{F_0(1+F_0-s^2)}[\delta(\Delta\omega-sqv)$$

$$+\delta(\Delta\omega+sqv)]\,\frac{d\Omega}{4\pi}\,d\Delta\omega,$$

where

$$\theta(y) = \begin{cases} 1 & y>0,\\ 0 & y<0.\end{cases}$$

This result has a simple physical meaning. As

can easily be seen, the frequency spectrum consists of a central part $-qv < \Delta\omega < qv$ and two narrow lines at $\Delta\omega = \pm sqv$. The central part corresponds to the Doppler broadening of the main line. Comparison of Eq. (10.21) with Eq. (8.16) shows that the side lines are the satellites of the Rayleigh scattering, which appear because of the possibility of the propagation of zeroth sound ($sv = u$). The ratio between the intensities of the central part and the satellites in general depends on the scattering angle. In the two extreme cases, however, of high temperatures ($T \gg \hbar\omega u/c$) and low temperatures ($T \ll \hbar\omega u/c$) this ratio does not depend on the angle. One can carry out numerical computations for He^3, using unknown (sic) parameters. It then turns out that for the high-temperature case about 20 percent of the whole intensity is in the central part, and about 40 percent in each of the side lines. In the low-temperature case, owing to the quantum factor, the distribution is cut off on the side of positive $\Delta\omega$. In particular, the only one of the two satellites left is the Stokes line with $\Delta\omega = -uq$, which gets 90 percent of the intensity. In this case the central part receives only 10 percent of the whole intensity.

The total scattered intensity is obtained by integrating Eq. (10.23) over $d\Delta\omega$ and $d\Omega$. For the high-temperature case

$$\left(T \gg \hbar\Delta\omega \sim \hbar\omega\frac{u}{c}\right)$$

it is

$$h = \frac{\omega^4 T}{6\pi c^4}\left(\frac{\partial D}{\partial N}\right)^2\left(\frac{d\tau_p}{d\varepsilon}\right)_{\varepsilon=\mu} J_1, \qquad (10.24)$$

where J_1 is a numerical integral, which in the case of He^3 is about 0.5. In the quantum limiting case ($T \ll \hbar\omega u/c$) we have:

$$h = \frac{\hbar\omega^5 v}{6\pi c^4}\left(\frac{\partial D}{\partial N}\right)\left(\frac{d\tau_p}{d\varepsilon}\right)_{\varepsilon=\mu} J_2. \qquad (10.25)$$

Here J_2 is a numerical integral, which in the case of He^3 is about 0.2. To carry out quantitative calculations from these expressions for the case of He^3, we need to know the value of $\partial D/\partial N_0$. There are no measurements of this quantity; therefore we take $D - 1$ to be proportional to N, and use the proportionality factor as given by the data on liquid He^4. For wavelength $\lambda = 5461$ A the index of refraction of liquid He^4 is 1.027. This gives $\partial D/\partial N = 2.5 \times 10^{-24}$. Substitution in Eqs. (10.24) and (10.25) gives (T in °K):

$$h(He^3)\ 10^{-69}\omega^4 T\ \text{cm}^{-1} \quad \text{for} \quad \omega \ll 2\cdot 10^{17} T\ \text{sec}^{-1},$$

$$h(He^3)\ 10^{-87}\omega^5\ \text{cm}^{-1} \quad \text{for} \quad \omega \gg 2\cdot 10^{17} T\ \text{sec}^{-1}.$$

It must be remembered that the frequency must satisfy the relation (10.1) i.e., $\Delta\omega \gg 1/\tau$, or

$$\omega \gg 10^{18} T^2\ \text{sec}^{-1}.$$

If this condition is not fulfilled, the line width will be too large. Thus in the visible range of frequencies we need temperatures below 0.05°K. It is not hard to see that at temperatures of the order of 0.01° and frequencies in the visible region we have $h(He^3) \sim 10^{-9}\text{cm}^{-1}$, which is of course too small for the effect to be measurable.* But owing to the fact that ω occurs raised to a high power in the expression for h, it is possible that one may be able to measure the scattering in the ultraviolet region.

APPENDICES

MICROSCOPIC THEORY OF THE FERMI FLUID

A 1. The Rarefied Fermi Gas

In this section we consider the properties of a non-ideal Fermi gas in which the dimensions of the particles are small in comparison with the mean free path.[16] This model, a number of whose properties have been examined recently by Yang and Huang,[17] and also by Yang and Lee,[18] was called by those writers the hard-sphere model. In reality it possesses a more general character, and corresponds to a system of Fermi particles with arbitrary short-range forces having a radius of action small in comparison with the mean wavelength.

Such a model makes it possible to obtain expressions for the various quantities describing the Fermi fluid, for example the energy, the effective mass, and the function f, in the form of expansions in powers of $(a/\lambda\!\!\!^{-})$ (a is the radius and $\lambda\!\!\!^{-}$ is the wavelength). We shall carry out the calculation to within terms of the order $(a/\lambda\!\!\!^{-})^2$. By a procedure similar in principle one could determine also terms of several higher orders, but this is of no special interest. We begin with the calculation of the energy.

We use perturbation theory to take into account the interaction energy of the particles, which we write in the form (with the volume of the gas taken as unity for simplicity)

$$V = 2U \sum_{\substack{n_1 n_2 m_1 m_2 \\ n_1 < n_2}} a^+_{m_1} a^+_{m_2} a_{n_2} a_{n_1}, \qquad (A1.1)$$

where a_i^+ and a_i are creation and annihilation

*For liquid He^4 in the visible region $h \sim 10^{-8}\,\text{cm}^{-1}\psi^{15}$ (about the same value should hold for He^3 for $\Delta\omega\tau \ll 1$). Although measurements of h for He^4 have been carried out, they are at the limit of what is experimentally possible.

operators for the particles. The summation is taken with the restriction that the total momentum is conserved; furthermore the spin component for the state m_1 is equal to that for the state n_1, and similarly for m_2 and n_2. The position of U outside the sign of summation corresponds to the fact that the interaction between all pairs of particles is the same, and the scattering amplitude does not depend on the angle. In first approximation the quantity U is connected with this amplitude by the relation

$$U=\frac{4\pi a\hbar^2}{m} \qquad \text{(A1.2)}$$

(a is the scattering amplitude).

The first-order correction to the energy is the diagonal matrix element of V:

$$E_n^{(1)}=2U\sum_{n_1<n_2} N_{n_1}N_{n_2}Q_{n_1n_2}, \qquad \text{(A1.3)}$$

where the N_i are occupation numbers.

The factor Q_{ik} in Eq. (A1.3) takes account of the fact that Fermi particles with a scattering amplitude independent of the angle do not interact in the case of parallel spins. Thus we shall take

$$Q_{ik}=\frac{1}{4}-\sigma_i\sigma_k \qquad \text{(A1.4)}$$

(σ_i is the spin operator of the i-th particle).

Substituting Eqs. (A1.2) and (A1.4) in Eq. (A1.3), we get

$$E^{(1)}=(2\pi a\hbar^2/m)\,N^2/2. \qquad \text{(A1.5)}$$

To find the second-order correction we use the expression from perturbation theory

$$E_n^{(2)}=\sum_{m\neq n}|V_{nm}|^2/(E_n-E_m). \qquad \text{(A1.6)}$$

Substituting the expression (A1.1) in this formula, we get the following sum:

$$4U^2\sum_{\substack{n_1n_2m_1m_2\\ n_1<n_2}}\frac{N_{n_1}N_{n_2}(1-N_{m_1})(1-N_{m_2})Q_{n_1n_2}}{E_{n_1}+E_{n_2}-E_{m_1}-E_{m_2}}, \qquad \text{(A1.7')}$$

where N_i are the equilibrium occupation numbers and E_i the energies of the particles.

Because our purpose is to get the expansion of the energy in powers of a, we have to take into account the fact that the relation (A1.2) between U and the scattering amplitude is not exact, but is valid only to first order. When second-order terms are included we get instead of (A1.2) the following relations:

$$2U+4U^2\sum_{m_1m_2}Q_{n_1n_2}/(E_{n_1}+E_{n_2}-E_{m_1}-E_{m_2})=\frac{8\pi a\hbar^2}{m}. \qquad \text{(A1.2')}$$

If we get from this the expression of U in terms of a and substitute it in Eq. (A1.3), the resulting expression for $E^{(1)}$ contains terms proportional to a^2, which must naturally be assigned to the second-order correction. Taking this into account, we get the following value of the second-order correction to the energy:

$$E^{(2)}=2U^2\sum_{n_1n_2m_1m_2}\left\{\frac{N_{n_1}N_{n_2}(1-N_{m_1})(1-N_{m_2})\cdot Q_{n_1n_2}Q_{n_2}}{E_{n_1}+E_{n_2}-E_{m_1}-E_{m_2}}\right.$$
$$\left.-\frac{N_{n_1}N_{n_2}Q_{n_1n_2}}{E_{n_1}+E_{n_2}-E_{m_1}-E_{m_2}}\right\}. \qquad \text{(A1.7)}$$

In view of the fact that the expression in brackets is symmetric in n_1 and n_2, we replace the restriction $n_1 > n_2$ by the factor 1/2.

The meaning of the operation just performed lies in the fact that in actuality an expansion in powers of U is not what is used. The existence of a constant U would simply lead to an infinity in the energy, as can be seen directly from the formula (A1.7). In the present case the essential point is that the scattering amplitude a has a finite, and moreover a small, value, which makes possible an expansion in powers of this quantity.

In the first term of the expression (A1.7) the term with the product of four N_i's is equal to zero because of the fact that the denominator is antisymmetric under the interchange $n_1n_2 \rightleftarrows m_1m_2$, while its numerator is symmetric, and all the ranges of summation are identical. The two surviving terms with products of three N_i's are identical. Thus we get finally:

$$E^2=-4U^2\sum_{n_1n_2m_1m_2}\frac{N_{n_1}N_{n_2}N_{m_1}Q_{n_1n_2}}{E_{n_1}+E_{n_2}-E_{m_1}-E_{m_2}}. \qquad \text{(A1.8)}$$

This is the expression for the energy that is valid on the basic assumption $(a/\lambda) \ll 1$.

Our purpose is to obtain the characteristics of the degenerate Fermi gas. From the expression (A1.8) we find the energy of the ground state

$$E^{(2)}=-\frac{4U^2}{(2\pi\hbar)^3}\int_{|p_1|<p_0;}dp_1\int_{|p_2|<p_0;}dp_2\int_{|p_3|<p_0;}dp_3\int dp_4\,\frac{\delta(p_1+p_2-p_3-p_4)}{(p_1^2+p_2^2-p_3^2-p_4^2)\frac{1}{2m}}, \qquad \text{(A1.9)}$$

where p_0 is the limiting momentum.

According to Sec. 1 the energy of the excitations is given by the relation

$$\varepsilon_i=\frac{\delta E}{\delta N_i}. \qquad \text{(A1.10)}$$

Variation of the expressions (1.3) and (1.8) with respect to N_i gives

$$\varepsilon(p)=\frac{p^2}{2m}+\frac{UN}{2}+\frac{2U^2}{(2\pi\hbar)^9}\int dp_1\int dp_2\int dp_3 \qquad \text{(A1.11)}$$
$$\times\left[\frac{\delta(p_1+p_2-p-p_3)}{(p^2+p_3^2-p_1^2-p_2^2)/2m}-2\frac{\delta(p_1+p-p_2-p_3)}{(p^2+p_1^2-p_2^2-p_3^2)/2m}\right].$$

Thus the problem of calculating the energy of the ground state and the effective mass of the excitations reduces to the calculation of the integrals (A1.9) and (A1.11). The integration is rather cumbersome because of the high multiplicity of the integrals and the awkwardness of the regions of integration.

Instead of this we can use a simpler method, based on the introduction of the function f. If we introduce this function

$$f_{ik}=\delta^2 E/\delta N_i \delta N_k, \qquad \text{(A1.12)}$$

then according to Eqs. (2.5) and (8.7) we can determine the effective mass and the speed of low-frequency sound. By a suitable integration we can find the energy of the ground state from the speed of sound.

Thus the problem reduces to the determination of the quantity f. Taking the variations of the expressions (A1.3) and (A1.8) with respect to N_1, and then with respect to N_k, we find the following expression:

$$f=2UQ_{\sigma,\sigma'}-\frac{8U^2}{(2\pi\hbar)^3}\int dp_1 \int\limits_{|p_1|<p_0} dp_2 \times\left[Q_{\sigma\sigma'}\frac{\delta(p+p'-p_1-p_2)}{(p^2+p'^2-p_1^2-p_2^2)/2m}\right. \qquad \text{(A1.13)}$$

$$\left.+\frac{1}{4}\frac{\delta(p+p_1-p'-p_2)}{(p^2+p_1^2-p'^2-p_2^2)/2m}+\frac{1}{4}\frac{\delta(p'+p_1-p-p_2)}{(p'^2+p_1^2-p^2-p_2^2)/2m}\right].$$

In the calculation we shall at once set $|p| = |p_1| = p_0$. The integration in the second term of f is considerably simpler than those in Eqs. (A1.9) and (A1.11). We get as the result:

$$f(\chi)=\frac{2\pi a\hbar^2}{m}\left[1+2\left(\frac{3}{\pi}\right)^{\frac{1}{3}}aN^{\frac{1}{3}}\left(2+\frac{\cos\chi}{2\sin\frac{\chi}{2}}\ln\frac{1+\sin\frac{\chi}{2}}{1-\sin\frac{\chi}{2}}\right)\right]$$

$$-\frac{8\pi a\hbar^2}{m}(\sigma_1\sigma_2)\left[1+2\left(\frac{3}{\pi}\right)^{\frac{1}{3}}aN^{\frac{1}{3}}\left(1-\frac{\sin\frac{\chi}{2}}{2}\right)\ln\frac{1+\sin\frac{\chi}{2}}{1-\sin\frac{\chi}{2}}\right]. \qquad \text{(A1.14)}$$

A peculiarity of this expression deserves attention. For angles χ close to π the function f has a logarithmic singularity,

$$f(\chi)\sim\left(\frac{1}{4}-\sigma_1\sigma_2\right)\ln\frac{1}{\pi-\chi}. \qquad \text{(A1.15)}$$

It is clear that in this case the approximation we have used is, strictly speaking, not legitimate. The singularity of the function f at $\chi = \pi$ is a manifestation of the singularity in the scattering amplitude of excitations colliding at the angle π. The corresponding expression, obtained by summing the main terms of the perturbation theory, is proportional to

$$\left[1+2\left(\frac{3}{\pi}\right)^{\frac{1}{3}}aN^{\frac{1}{3}}\ln\frac{p_0^2}{\varepsilon}\right]^{-1}, \qquad \text{(A1.16)}$$

where $\epsilon = p^2 + p'^2 - 2p_0^2$.

In the case of positive a this expression goes to zero for $p^2 = p'^2 = p_0^2$.

If, however, $a < 0$ (this is possible for Fermi systems), the scattering amplitude has a pole near the Fermi surface. This corresponds to the possibility of the formation of bound pairs of excitations with opposite momenta, which was pointed out by Cooper[19] and is the main cause of the phenomenon of superconductivity in metals.[20]

Thus the expression found for f is not valid for angles close to π. But in view of the fact that the singularity is only logarithmic, it manifests itself only in the immediate neighborhood of the singular point. And since only integrals of f together with regular functions occur in the expressions with which we are concerned, the logarithmic singularity of the function f is of no importance.

Substituting Eq. (A1.4) in Eq. (2.5), we find the value of the effective mass

$$m/m^*=1-(8/15)(3/\pi)^{\frac{2}{3}}(7\ln 2-1)\,a^2N^{\frac{2}{3}}. \qquad \text{(A1.17)}$$

We note that if we insert here the value $m^* = 1.43$ m, which corresponds to liquid He^3, the resulting value of a is 1.6×10^{-8}; that is, it is of the order of magnitude of the gas-kinetic value of the diameter of the helium atom. Such a comparision of course has no meaning in a strict sense. The model under consideration cannot describe liquid He^3. This can already be seen from the fact that the quantity $(m^*-m)/m^*$, which in the theory should be a quantity of second order in a, is 1/3 for He^3.

Substituting the formula for f in the expression for the speed of sound, we get:

$$c^2=\frac{\pi^{\frac{4}{3}}}{3^{\frac{1}{3}}}N^{\frac{2}{3}}\frac{\hbar^2}{m^2}+2\frac{\pi a\hbar^2}{m^2}N\left[1+\frac{4}{15}\left(\frac{3}{\pi}\right)^{\frac{1}{3}}aN^{\frac{1}{3}}(11-2\ln 2)\right]. \qquad \text{(A1.18)}$$

From the value of c^2 just found it is not hard to obtain the energy of the ground state of the Fermi fluid. To do this we use the relation (8.2), $c^2 = N/m(\partial\mu/\partial N)$. This gives:

$$E=\int\mu\,dN=E^{(0)}+\frac{\pi a\hbar^2}{m}N^2\left[1+\frac{6}{35}\left(\frac{3}{\pi}\right)^{\frac{1}{3}}aN^{\frac{1}{3}}(11-2\ln 2)\right]. \qquad \text{(A1.19)}$$

Equation (A1.2) agrees with the result of Lee and Yang.[15]

The same result can also be obtained by direct integration in Eq. (A1.9). To calculate the integral in Eq. (A1.9) it is convenient to introduce new variables

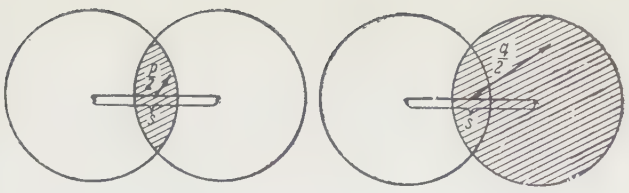


FIG. 2

$$\mathbf{p}=\mathbf{p}_1-\mathbf{p}_2;\quad \mathbf{q}=\mathbf{p}_3-\mathbf{p}_4;\quad \mathbf{s}=\mathbf{p}_1+\mathbf{p}_2=\mathbf{p}_3+\mathbf{p}_4.$$

In these variables $E^{(2)}$ takes the form

$$E^{(2)}=-\frac{ml^2}{4(2\pi\hbar)^2}\int ds\int dp\int dq\,\frac{1}{p^2-q^2},$$

where the region of integration for the vector s is $0 < |s| < 2p_0$, and the regions of values for the vectors p and q are shown in Fig. 2. If we introduce the variables $x_q = \cos(\mathbf{q}, \mathbf{s})$ and $x_p = \cos(\mathbf{p}, \mathbf{s})$, $E^{(2)}$ is given by

$$E^{(2)}=-\frac{8mU^2\pi^3}{(2\pi\hbar)^9}\int_0^{2p_0} s^2\,ds\int_0^1 dx_p\int_0^{z(x_p)} p^2\,dp\int_{-1}^1 dx_q\int_0^{z(x_q)} q^2\,dq\,\frac{1}{p^2-q^2},$$

where z(x) satisfies the relation

$$z^2+2zsx+s^2=4p_0^2.$$

By a series of transformations and integrations by parts with respect to dx_p and dx_q we then find:

$$E^2=\frac{2mU^2p_0^7}{\pi^6\hbar^9}\int_0^1 s^2\,ds\left[\int_0^{1+s} p^2\,dp\int_0^{1-s} q^2\,dq\right.$$

$$\left.+\frac{1}{4s^2}\int_{\sqrt{1-s^2}}^{1+s} p\,dp\,(1-p^2-s^2)\int_{\sqrt{1-s^2}}^{1-s} q\,dq\,(1-q^2-s^2)\right]\frac{1}{p^2-q^2}.$$

Proceeding further to integrate by parts with respect to ds and then performing the remaining integration, we get:

$$E^{(2)}=(6/35)(3/\pi)^{\frac{1}{3}}(11-2\ln 2)\,aN^{\frac{1}{3}}E^{(1)}.$$

Here we have inserted the expression (A1.2) for U and have set $p_0 = \hbar(3\pi^2N)^{1/3}$. This result is identical with the second-order term in Eq. (A1.19).

A.2. Microscopic Theory of the Fermi Fluid for T = 0

Here we present the general microscopic theory of the Fermi fluid at T = 0, as developed recently by Landau;[18] this theory is valid for arbitrary interactions. As we shall see, this treatment provides a basis for the phenomenological theory and an understanding of the physical meaning of the function f that appears in it.

The microscopic treatment is based on methods taken from the quantum theory of fields. As is well known, in quantum field theory one works with vacuum expectation values of chronological products of field operators (i.e., products of operators taken in the order of succession of their times). These quantities are the so-called Green's functions. In our case, instead of vacuum expectation values we shall consider expectation values for the ground state of the system. It is not hard to verify that such expectation values have the same properties as the Green's functions of the field theory, and in particular that in principle they can be calculated by means of the Feynman-diagram technique.

For definiteness we shall first give the name of Green's function to the quantity

$$G_{1,2}=-i\langle T(\psi_1\psi_2^+)\rangle, \qquad \text{(A2.1)}$$

where the brackets $\langle\ \rangle$ denote the expectation value for the ground state, ψ and ψ^+ are the field operators of the Fermi particles, the indices 1, 2 denote sets of space, time, and spin coordinates, and T, as usual, denotes the time-ordered product with change of sign when ψ_1 and ψ_2^+ are interchanged. In the absence of external fields the Green's function depends only on the differences of the coordinates and the times. As for the spin indices, if the interaction of the particles does not depend on their spins, then $G_{\alpha\beta}\sim\delta_{\alpha\beta}$.

We denote the Fourier transform of the Green's function with respect to the coordinate and time differences by $G(p)\,\delta_{\alpha\beta}$ [here p is the four-dimensional vector $(\mathbf{p}, \epsilon)$]. The poles of G(p) determine the energies of the quasi-particles. If the magnitude of p is close to p_0 and ϵ is close to the limiting energy μ, then G(p) has the form

$$G(p)=\frac{a}{s-\mu-v(p-p_0)+i\delta}. \qquad \text{(A2.2)}$$

Here a is a positive constant, and δ is a small quantity which agrees in sign with the sign of $\epsilon-\mu$, or, close to a pole, with the sign of $p-p_0$ (δ is introduced to fix the rule for going around the poles, so that it will correspond to taking the T-product in the x, t representation).

Let us now introduce the expectation value of the time-ordered product of four operators

$$\Phi_{12,34}=\langle T(\psi_1\psi_2\psi_3^+\psi_4^+)\rangle. \qquad \text{(A2.3)}$$

We give the name of vertex part $\Gamma_{12,34}$ to the quantity defined by the following relation:

$$\Phi_{12,34}=G_{1,3}\,G_{2,4}-G_{1,4}\,G_{2,3}$$

$$+i\sum_{1',2',3',4'}G_{1,1'}G_{2,2'}G_{3,3'}G_{4,4'}\Gamma_{1'2',3'4'}. \qquad \text{(A2.4)}$$

The Fourier transform of the vertex part has the form

$$(2\pi\hbar)^4\,\Gamma_{\alpha\beta,\,\gamma\delta}(p_1p_2,\,p_3p_4)\,\delta(p_1+p_2-p_3-p_4).$$

It is clear that the function $\Gamma_{\alpha\beta\ \gamma\delta}(P_1P_2, P_3P_4)$ changes sign on the interchange $\alpha p_1 \rightleftarrows \beta p_2$. As in the quantum field theory, the function Γ with a definite relation between the components of each of the vectors P [for example, near the Fermi surface $\epsilon - \mu = v(p - p_0)$], multiplied by a^2, plays the role of the scattering amplitude of the quasi-particles.

Let us consider Γ in the case in which p_1 is close to p_3, and consequently p_2 is close to p_4, and introduce the definition

$$\Gamma(p_1p_2,\ p_1+K\ p_2-K)\equiv\Gamma(P_1P_2,\ K),$$

where $K = (\mathbf{k}, \omega)$ is a small four-vector.

If it were possible to calculate such a vertex part by perturbation theory, then in first order one would have to take the diagrams shown in Figs. 3, I. The corresponding expressions contain integrals of products of two Green's functions. Whereas for diagrams (a) and (b) the case $K = 0$ is not in any way singled out, in case (c) the poles of the two Green's functions approach each other for $K = 0$, which, as we shall see, leads to the appearance of a singularity of Γ.

If we denote by $\Gamma^{(1)}$ the set of all possible diagrams for Γ that do not have "singular elements" [i.e., integrals of G(p)G(p + K)], it is not hard to see that the total Γ is obtained by summation of the "ladders" shown in Fig. 3, II. This can be accomplished by solving the integral equation

$$\Gamma_{\alpha\beta,\ \gamma\delta}(P_1P_2;\,K)=\Gamma^{(1)}_{\alpha\beta,\ \gamma\delta}(P_1P \tag{A2.5}$$

$$\frac{i}{(2\pi\hbar)^4}\int\Gamma^{(1)}_{\alpha\varepsilon,\ \gamma\zeta}(P_1Q)\,G(Q)\,G(Q+K)\,\Gamma_{\zeta\beta,\ \varepsilon\delta}(QP_2;\,K)\,d^4Q.$$

Here, in view of the fact that $\Gamma^{(1)}$ has no singularity at $K = 0$, we simply set $K = 0$ everywhere in $\Gamma^{(1)}$.

Let us now examine the product in the integrand. The integral over Q will consist of the principal value and a term arising from the passage around the poles. Because the arguments of the two G functions are nearly equal, we can assume that all the other parts of the integral have the same values at these poles. The passage around the poles then gives a contribution only in the case in which the detours are on opposite sides of the real axis. This means either $q < p_0$, $|\mathbf{q} + \mathbf{k}| > p_0$, or vice versa. Because of the smallness of $\mathbf{k}$ it is not hard to see that here $q \approx p_0$ and $\epsilon_q \approx \mu$. Thus in the part of the integral over Q which comes from the passage around the poles the product G(Q)G(Q + K) can be replaced by $A\,\delta(\epsilon - \mu)\,\delta(q - p_0)$.

The coefficient A can be found by integrating G(Q)G(Q + K), and is found to be $(2\pi i a^2 k\cos\theta)/(\omega - vk\cos\theta)$, where θ is the angle between $\mathbf{q}$ and $\mathbf{k}$. Thus the product G(Q)G(Q + K) can be written in the form

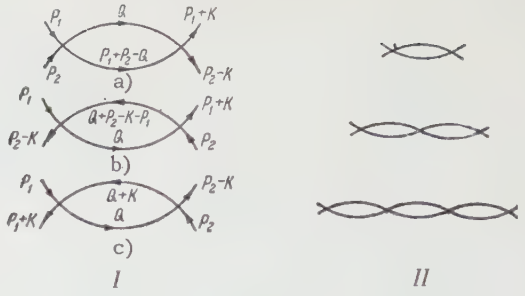


FIG. 3

$$G(Q)\,G(Q+K)=2\pi i a^2\frac{k\cos\vartheta}{\omega-vk\cos\vartheta}\,\delta(\varepsilon-\mu)\,\delta(q-p_0)+\varphi(Q), \tag{A2.6}$$

where φ corresponds to the principal value of the integral and has no singularities (therefore one puts $K = 0$ in this term).

If we consider the expression (A2.6) in the limit $K = 0$, then it makes a great difference how this limit is taken, i.e., what the ratio of ω and k approaches. The same is true about Γ in the limit $\omega \to 0$, $\mathbf{k} \to 0$.

First of all let us consider Γ in the limit $\mathbf{k} \to 0$, $(k/\omega) \to 0$. For this limit, which we denote by Γ^{ω}, we get from Eqs. (A2.5) and (A2.6):

$$\Gamma^{\omega}_{\alpha\beta,\ \gamma\delta}(P_1P_2)=\Gamma^{(1)}_{\alpha\beta,\ \gamma\delta}(P_1P_2) \tag{A2.7}$$

$$-\frac{i}{(2\pi\hbar)^4}\int\Gamma^{(1)}_{\alpha\xi,\ \gamma\zeta}(P,\,Q)\,\varphi(Q)\,\Gamma^{(\omega)}_{\zeta\beta,\ \varepsilon\delta}(QP_2)\,d^4Q.$$

It is not hard to eliminate $\Gamma^{(1)}$ from the two equations (A2.5) and (A2.7). Using Eq. (A2.6), we find after simple transformations

$$\Gamma_{\alpha\beta,\ \gamma\delta}(P_1P_2;\,K)=\Gamma^{(\omega)}_{\alpha\beta,\ \gamma\delta}(P_1P_2) \tag{A2.8}$$

$$+\frac{a^2p_0^2}{(2\pi\hbar)^3}\int\Gamma^{\omega}_{\alpha\varepsilon,\ \gamma\zeta}(P,\,Q)\,\Gamma_{\zeta\beta,\ \varepsilon\delta}(QP_2,\,K)\,\frac{k\cos\vartheta}{\omega-vk\cos\vartheta}\,d\Omega.$$

Let us now take the other limit, namely $K \to 0$, $(\omega/k) \to 0$. This limiting value, multiplied by a^2, corresponds physically to the forward scattering amplitude for quasi-particles with energy on the Fermi surface. We denote this quantity by Γ^{k}. From Eq. (A2.8) we find the relation between $a^2\Gamma^{k}$ and $a^2\Gamma^{\omega}$

$$a^2\Gamma^{k}_{\alpha\beta,\ \gamma\delta}(P_1P_2)=a^2\Gamma^{\omega}_{\alpha\beta,\ \gamma\delta}(P_1P_2) \tag{A2.9}$$

$$-\frac{p_0^2}{v(2\pi\hbar)^3}\int a^2\Gamma^{\omega}_{\alpha\varepsilon,\ \gamma\zeta}(P,\,Q)\,a^2\Gamma^{k}_{\zeta\beta,\ \varepsilon\delta}(QP_2)\,d\Omega.$$

Let us now study the poles of the function $\Gamma(P_1P_2; K)$. Just as the poles of the function G(P) give the relation between the energies and momenta of the Fermi excitations, the poles of the function Γ characterize the "two-particle" excitations, in

other words the Bose excitations. In view of the fact that in the neighborhood of a pole $\Gamma(P_1 P_2; K) \gg \Gamma^{\omega}(P_1 P_2)$, we can neglect the term Γ^{ω} in the right member. Furthermore it must be noted that the variable P_2, and also the indices β and δ play the part of parameters in the equation. The function Γ can be represented in the form of a product $\chi_{\alpha\beta}(P_1; K)\,\chi_{\beta\delta}(P_2)$, and $\chi(P_2)$ cancels on the two sides of the equation. If we also introduce the notation

$$\nu_{\alpha\gamma}(\mathbf{n}) = \frac{\mathbf{n}\cdot\mathbf{k}}{\omega - v\mathbf{n}\cdot\mathbf{k}}\chi_{\alpha\gamma}(P_1; K),$$

where $\mathbf{n}$ is the unit vector in the direction of P_1, we get for $\nu_{\alpha\gamma}$ the equation

$$(\omega - v\mathbf{n}\cdot\mathbf{k})\nu_{\alpha\gamma}(\mathbf{n}) = \mathbf{n}\cdot\mathbf{k}\,\frac{p_0^2 a^2}{(2\pi\hbar)^3}\int \Gamma^{\omega}_{\alpha\varepsilon,\,\gamma\zeta}(\mathbf{n},\mathbf{l})\,\nu_{\zeta\varepsilon}(\mathbf{l})\,d\Omega. \quad \text{(A2.10)}$$

Comparing this equation with Eqs. (8.14) and (10.4), we see without difficulty that it is the same as the equation for the zeroth sound and the spin waves, with the quantity $a^2\Gamma^{\omega}$ playing the role of the function f. This confirms the equation for the zeroth sound obtained in the phenomenological theory, and thus confirms the hypothesis made in that theory about the functional dependence of the energy of the excitations on the distribution function. On the other hand, owing to Eq. (A2.9), we have established a connection between the function f and the scattering amplitude of the quasi-particles at the angle 0°.

Let us examine this relation in greater detail. Denoting the amplitude multiplied by $(d\tau/d\epsilon)_{\epsilon=\mu}$ by $A(\mathbf{n}_1\cdot\sigma_1, \mathbf{n}_2\cdot\sigma_2)$, we get:

$$A(\mathbf{n}_1\cdot\sigma_1, \mathbf{n}_2\cdot\sigma_2) = \Phi(\mathbf{n}_1\cdot\sigma_1, \mathbf{n}_2\cdot\sigma_2) - \frac{1}{2}\mathrm{Sp}_{\sigma'}\int \Phi(\mathbf{n}_1\cdot\sigma_1, \mathbf{n}'\cdot\sigma')A(\mathbf{n}'\cdot\sigma', \mathbf{n}_2\cdot\sigma_2)\,\frac{d\Omega}{4\pi}. \quad \text{(A2.11)}$$

As regards their spin dependence, Φ and A each contain two terms: one independent of the spins, and one proportional to $\sigma_1\cdot\sigma_2$. If we write

$$A(\mathbf{n}_1\cdot\sigma_1, \mathbf{n}_2\cdot\sigma_2) = B(\mathbf{n}_1\cdot\mathbf{n}) + \sigma_1\cdot\sigma_2\cdot C(\mathbf{n}_1, \mathbf{n}_2), \quad \text{(A2.12)}$$

B and C satisfy the equations

$$B(\mathbf{n}_1, \mathbf{n}_2) = F(\mathbf{n}_1, \mathbf{n}_2) - \int F(\mathbf{n}_1, \mathbf{n}')\,B(\mathbf{n}', \mathbf{n}_2)\frac{d\Omega}{4\pi}, \quad \text{(A2.13)}$$

$$C(\mathbf{n}_1, \mathbf{n}_2) = Z(\mathbf{n}_1, \mathbf{n}_2) - \frac{1}{4}\int Z(\mathbf{n}_1, \mathbf{n}')\,C(\mathbf{n}', \mathbf{n}_2)\frac{d\Omega}{4\pi}.$$

Being scalars, the quantities B and C, and also F and Z, can depend only on cos χ. If we introduce expansions in spherical harmonics, $B(\chi) = \Sigma B_l P_l(\cos\chi)$, and so on, then we can easily get relations between the coefficients B_l and F_l, and between C_l and Z_l:

$$B_l = \frac{F_l}{1 + \frac{F_l}{2l+1}}, \quad C_l = \frac{Z_l}{1 + \frac{Z_l}{4(2l+1)}}. \quad \text{(A2.14)}$$

For example, let us find the scattering amplitude of the excitations for a rarefied Fermi gas. From Eqs. (A1.16) and (A2.13) we get:

$$A(\chi, \sigma_1\cdot\sigma_2) = 8\left(\frac{3}{\pi}\right)^{\frac{1}{3}} aN^{\frac{1}{3}}\left(\frac{1}{4} - \sigma_1\cdot\sigma_2\right) \quad \text{(A2.15)}$$

$$+ 16\left[\left(\frac{3}{\pi}\right)^{\frac{1}{3}} aN^{\frac{1}{3}}\right]^2\left[\frac{1}{4} - \sigma_1\cdot\sigma_2 - \frac{\cos\chi}{8\sin\frac{\chi}{2}} - \ln\frac{1-\sin\left(\frac{\chi}{2}\right)}{1+\sin\left(\frac{\chi}{2}\right)}\right.$$

$$\left. - \sigma_1\cdot\sigma_2\left(1 - \frac{\sin\left(\frac{\chi}{2}\right)}{2}\ln\frac{1-\sin\left(\frac{\chi}{2}\right)}{1+\sin\left(\frac{\chi}{2}\right)}\right)\right].$$

We note that in the case of parallel spins this expression goes to zero for $\chi = 0$, as it must according to the Pauli principle (for f this does not occur).

[1] L. D. Landau, J. Exptl. Theoret. Phys. (U.S.S.R.) **30**, 1058 (1956), Soviet Phys. JETP **3**, 920 (1956).

[2] I. M. Khalatnikov and A. A. Abrikosov, J. Exptl. Theoret. Phys. (U.S.S.R.) **32**, 915 (1957), Soviet Phys. JETP **5**, 745 (1957).

[3] Abraham, Osborne, and Weinstock, Phys. Rev. **98**, 551 (1955).

[4] Fairbank, Ard, and Walters, Phys. Rev. **95**, 567 (1954).

[5] A. A. Abrikosov and I. M. Khalatnikov, J. Exptl. Theoret. Phys. (U.S.S.R.) **32**, 1083 (1957), Soviet Phys. JETP **5**, 887 (1957).

[6] I. Ia. Pomeranchuk, J. Exptl. Theoret. Phys. (U.S S.R.) **20**, 919 (1950).

[7] K. N. Zinov'eva, J. Exptl. Theoret. Phys. (U.S.S.R.) **34**, 609 (1958), Soviet Phys. JETP **7**, 421 (1958).

[8] L. D. Landau, J. Exptl. Theoret. Phys. (U.S.S.R.) **32**, 59 (1957), Soviet Phys. JETP **5**, 101 (1957).

[9] Lagner, Sydoriak, and Roberts, Program of the Fifth International Conference on Low-Temperature Physics and Chemistry, p. 41. University of Wisconsin 1957.

[10] I. M. Khalatnikov and A. A. Abrikosov, J. Exptl. Theoret. Phys. (U.S.S.R.) **33**, 110 (1957), Soviet Phys. JETP **6**, 84 (1958).

[11] A. A. Abrikosov and I. M. Khalatnikov, J. Exptl. Theoret. Phys. (U.S.S.R.) **34**, 198 (1958), Soviet Phys. JETP **7**, 135 (1958).

[12] L. D. Landau and E. M. Lifshitz, Электродинамика сплошных сред (Electrodynamics of Continuous Media), Gostekhizdat, Moscow 1957.

[13] S. M. Rytov. Теория электрических флуктуаций

и теплового излучения (Theory of Electrical Fluctuations and Thermal Radiation), Acad. Sci. Press, 1953.

[14]L. D. Landau and E. M. Lifshitz, J. Exptl. Theoret. Phys. (U.S.S.R.) **32**, 618 (1957), Soviet Phys. JETP **5**, 512 (1957).

[15]V. L. Ginzburg, J. Exptl. Theoret. Phys. (U.S.S.R.) **13**, 243 (1943).

[16]A. A. Abrikosov and I. M. Khalatnikov, J. Exptl. Theoret. Phys. (U.S.S.R.) **33**, 1154 (1957), Soviet Phys. JETP **6**, 888 (1958).

[17]K. Huang and C. N. Yang, Phys. Rev. **105**, 767 (1957).

[18]T. D. Lee and C. N. Yang, Phys. Rev. **105**, 1119 (1957).

[19]L. N. Cooper, Phys. Rev. **104**, 1189 (1956).

[20]Bardeen, Cooper, and Schrieffer, Phys. Rev. **106**, 162 (1957).

[21]L. D. Landau, J. Exptl. Theoret. Phys. (U.S.S.R.) (in press).

Translated by W. H. Furry

PHYSICO-CHEMICAL PHENOMENA IN THE DEFORMATION OF METALS

V. I. LIKHTMAN and E. D. SHCHUKIN

Usp. Fiz. Nauk **66**, 213-245 (October, 1958)

A new scientific field has arisen during the last ten years, primarily in the papers by the group of Soviet scientists headed by Academician P. A. Rebinder. This is physico-chemical mechanics, which borders on physical chemistry, molecular physics, and mechanics of materials. The basic aim of this new borderline field of knowledge[9] consists of (1) the clarification of the laws and the mechanism of formation of solids of a given structure and with prescribed mechanical properties, and (2) the study of the mechanism of the processes of deformation, working, and fracture of solids, taking into account the effect of physico-chemical factors such as the composition and the structure of the solid, the temperature, and the surrounding medium.

By now considerable experimental material has been accumulated in this field, which is of great significance for the theory of the formation and fracture of solids, and also for the determination of optimum conditions of working them. At the same time in recent years considerable advances have been made in the field of theoretical treatment of the newly discovered laws with the aid of the formalism of dislocation theory.

This article presents the basic results of investigations carried out recently in the field of physico-chemical mechanics of metals, with particular attention to the theoretical treatment of these results from the point of view of dislocation theory.

It has already been established for a long time, through the work of P. A. Rebinder and his associates, that the ability of solids to resist deformation and fracture is reduced as a result of adsorption from the surrounding medium. This phenomenon is associated with a decrease in the surface energy of the deformed solid. The effects of the reduction of strength due to adsorption and of the decrease in resistance to deformation are of a kinetic nature. As a result of a decrease in the surface tension of the solid over its surfaces, they are brought about by the greater ease of formation and development of various types of structural defects and of incipient shear, which in the course of deformation develop over the boundary with the surrounding medium.[1-6*]

The strength must, evidently, be decreased by adsorption indefinitely with decreasing interphase surface energy at the boundary between the solid and the surrounding medium, since the magnitude of the interphase surface energy determines the probability and the work of formation of surface defects (incipient shear or fracture) appearing in the process of deformation.

The largest adsorption effects due to the surrounding medium, as has been shown,[4,5] are observed during prolonged exposure to appropriate stresses, under conditions of creep or fatigue testing.

Organic surface-active substances (the usual active components of lubricants) can be used only at relatively low temperatures. Surface-active metals and alloys with comparatively low melting points are considerably more active with respect to metals and give rise to a large reduction in their surface tension, i.e., they are more strongly adsorbed.

The case of sufficiently low values of interphase surface tension σ_{12} at the boundary between two phases corresponds, according to Volmer's

*The papers by Benedicks and Sato[7] appeared long after P. A. Rebinder's first publications on this subject.[1] They pointed out (without mentioning the original source) essentially the same effect of reduction in the strength of solids under the action of wetting liquid media. However, this was accompanied by an erroneous conclusion with respect to the alleged possibility of determining the surface tension of a solid σ_1 by making use of Antonov's empirical rule (and moreover applying it erroneously to mutually nonsaturated phases) and extrapolating linearly (!) the initial decrease in strength P_m with increasing surface tension of the wetting liquid σ_2 down to the value $P_m = 0$, which should correspond to $\sigma_1 = \sigma_2$. This unjustified extrapolation leads to values of σ_1, for example in the case of steel, lower by an order of magnitude than reliable values computed by various indirect methods. In going over to large reductions of strength in liquids, with respect to which the given solid is more lyophilic, we would obtain concave curves which approach the horizontal axis ($P_m = 0$) at considerably higher values of σ_2. It is clear that when a contact angle is definitely not formed the work of spreading is positive and large, and Antonov's rule is obviously inapplicable.

conception,[8] to spontaneous dispersion down to particles of colloid size under the action of thermal motion. In this case a thermodynamically stable colloid solution (suspension or emulsion) is formed which is a completely dispersed two-phase system of two mutually saturated phases that are in equilibrium. The increase in surface energy in the case of such a dispersion is compensated by an increase in the entropy of the system as a result of the uniform distribution of the particles throughout the whole volume of the medium.

An approximate condition for spontaneous dispersion consists of the requirement that $\sigma_{12} < \sigma_m$, where $\sigma_m = \gamma kT/\delta_m^2$ (δ_m is the average size of a mosaic block, $\sim 10^{-6}$ cm), or, according to Rebinder, the dispersion criterion $Di = \sigma_{12}\delta_m^2/kT\gamma$ (where $\gamma \sim 30$[68]) must satisfy the condition $Di < 1$. Spontaneous dispersion can occur even at the relatively high values $\sigma_{12} \cong 0.2$ erg-cm^{-2}. As the temperature is increased σ_{12} decreases sharply because of the increase in mutual solubility. The limiting case of the critical point, $\sigma_{12} = 0$, corresponds according to Mendeleev to the disappearance of the colloid phase, i.e., to infinite mutual solubility with the formation of a true (single-phase) solution.

For large values of σ_{12} ($\sigma_{12} > \sigma_m$) the dispersion cannot be spontaneous, i.e., it occurs only under the action of external forces, but in this case the strength of a solid in a given liquid phase $(P_m)_A$ shows a greater reduction with diminishing σ_{12} in comparison with its strength in vacuo $(P_m)_0$. Under this condition the criterion for dispersion will be of the form:

$$Di = \frac{\sigma_{12}S_m - 3PV_m}{\gamma kT} < 1,$$

while the effect of reduction of strength may attain very considerable magnitude.

Until recently investigations of the effect of low-melting-point metallic fluxes on the mechanical properties of high-melting-point metals and alloys were carried out using polycrystalline samples tested under compression, bending, or tension. The most general result of these investigations is that in the presence of metallic fluxes the plasticity and the strength of samples deformed in such a medium is sharply reduced.[10-16,69] At the same time it was found that the magnitude of the effect of the reduction of plasticity and strength under the action of a flux depends to a large extent on the nature of the metal being deformed and of the low-melting-point flux surrounding it.

In the opinion of many authors, fluxes which form solid solutions or intermetallic compounds with the basic metal are more effective, while the absence of interaction between a metal and a flux leads to a considerable decrease in the magnitude of the effect. Moreover, the effect of flux action depends on the mechanical characteristics of the initial sample. The effect of the medium is more pronounced if the sample has greater strength and hardness. In the case of steel, brittle fracture can occur both in the elastic and in the plastic range, depending on the elastic limit of the steel.

An essential role is also played by the conditions under which the experiments are carried out: the nature of the stressed state, the rate of deformation or the manner in which the stress is applied, the time of contact of the sample with the slux, and the temperature at which the deformation is produced. Thus, for example, a flux does not produce any effect under the action of compressive stresses, and produces a large effect under the action of tensile stresses. The effectiveness increases as the rate of loading or the rate of deformation of the sample is decreased, or as the time of contact between the sample and the flux before the experiment is increased; the effectiveness of the flux increases appreciably as the temperature of the experiment is increased.

The extensive experimental material accumulated by now on the effect of metallic fluxes on the mechanical properties of high-melting-point metals has not yet received an exhaustive explanation. Many authors attribute these phenomena to the usual intercrystallite corrosion by assuming that as the flux diffuses along the grain boundaries it weakens the binding between the grains by means of some kind of a mechanism.

Of considerable importance for the correct understanding of the phenomenon of the decrease in the strength of metals under the influence of low melting point metallic fluxes are the papers of S. T. Kishkin and Ya. M. Potak, who emphasize the relation between these phenomena and the effect of the decrease in the strength of metals due to adsorption.[16]

We present below the main results of the experimental investigation of the effect of low-melting-point metallic fluxes on the mechanical properties of metals of higher melting points obtained for single crystals of zinc, tin and cadmium by L. A. Kochanova, L. S. Bryukhanova and V. A. Labzin of our laboratory.[17,18,19]

Tin, lead, alloys of tin with lead in varying concentrations, and mercury were used as surface-active fluxes. Single crystals of different orientation were grown by the method of zone crystalliza-

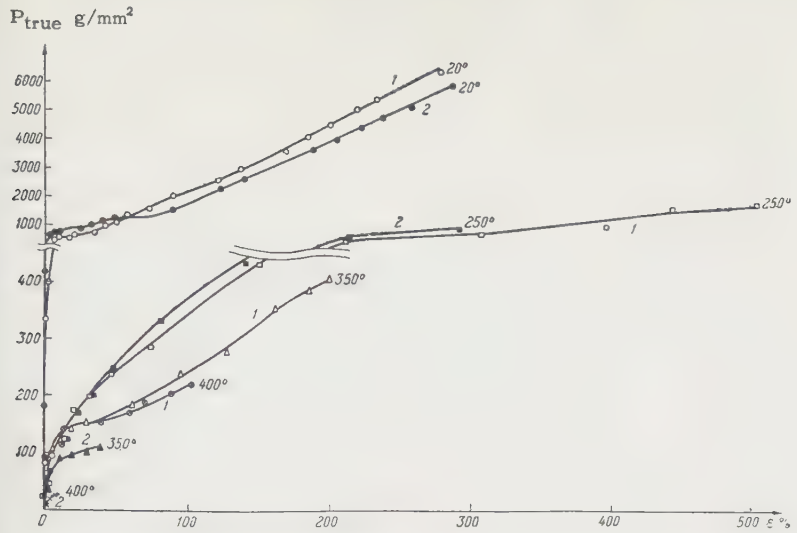


FIG. 1. Dependence of the true stress $P = P_0 (1 + \varepsilon/100)$ on the relative elongation ε of zinc single crystals ($\chi_0 = 44°$) without external coating (1) and coated with a film of tin (2). The temperature of the crystals is shown on the graph.

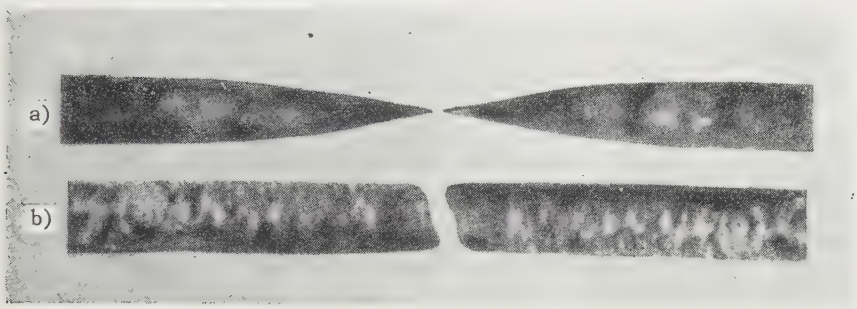


FIG. 2. The usual plastic fracture of zinc single crystals without external coating (a) and brittle fracture (b) of zinc single crystals coated with a film of tin. Both fractures took place at 400° C.

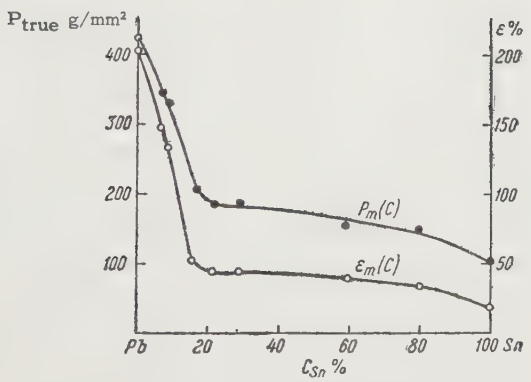


FIG. 3. Dependence of the true strength P_m and of the greatest elongation ε_m preceding fracture of zinc single crystals ($\chi_0 = 42°$) on the concentration of tin in an alloy of tin with lead at 350° C.

tion developed in our laboratory.[20]

Figure 1, in which the axes represent the true stress and the relative elongation, presents stress-strain curves for single crystals of zinc at different temperatures. The elongation took place at a constant deformation rate of 15% min^{-1}. The diameter of the single crystals was 0.5 mm. Black points indicate the curves for the elongation of crystals whose deformation was carried out in tin (a thin film of tin of thickness 2 to 5 μ was deposited electrolytically on the surface of the single crystal of zinc).

Below the melting point of the eutectic ZnSn, the effect of a thin film of tin on the surface of the zinc single crystal results only in a small increase of strength. At higher temperatures (350° and 400°C), in the presence of tin, the strength and plasticity decrease rather sharply (by a factor of 8 to 10). At the same time the nature of the fracture changes from ductile fracture in air to brittle fracture in an active medium (Fig. 2). The pronounced effect of the temperature (in going from 350° to 400°C) is evidently associated with an increase in the solubility of zinc in liquid tin (45% at 350° and 92% at 400°C), which corresponds to a sharp decrease in the interphase surface tension.

A similar investigation was made of the influence of lead-tin fluxes of different tin content on the stress-strain curve of single crystals of zinc. Molten lead taken by itself turned out to be relatively weakly surface-active with respect to zinc. Figure 3 shows that even relatively small additions to the flux of surface-active tin give rise to a fairly sharp decrease in the strength and in the plasticity of single crystals of zinc. At 20% this effect reaches a limit, and further reduction occurs only at a tin content of greater than 80% (when pure tin is approached), which is evidently associated with a sharp increase in the solubility of zinc in such fluxes.

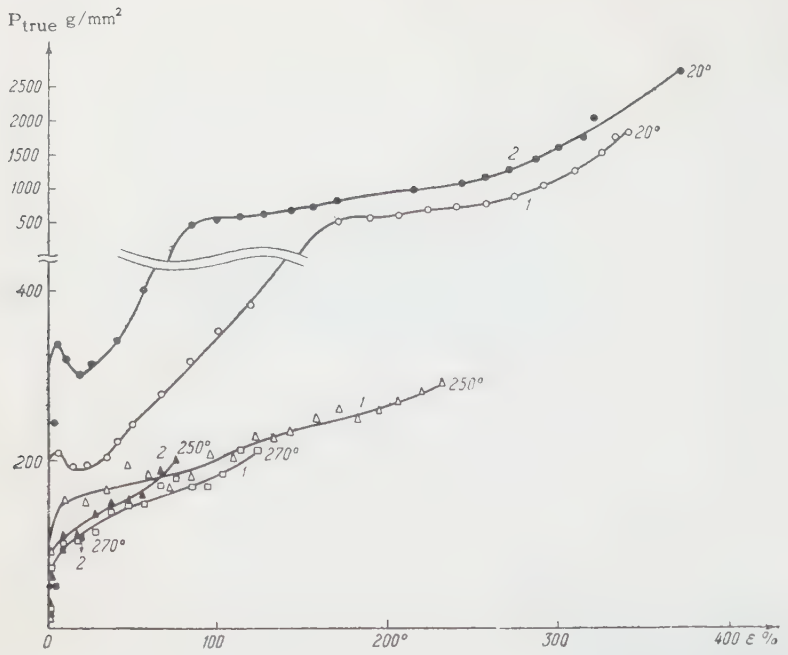


FIG. 4. Dependence of the true stress P on the relative elongation ε of cadmium single crystals ($\chi_0 = 39°$) without external coating (1) and coated with a thin film of tin (2). The temperature of the crystals is shown on the graph.

TABLE I. Rate of steady creep of zinc single crystals tested in an inert gas $(v_m)_0$ and in the presence of a thin film of tin on the surface of the single crystals $(v_m)_A$.

t °C	P (g/mm²)	$(v_m)_0$ (min⁻¹)	$(v_m)_A$ (min⁻¹)	$\frac{(v_m)_A}{(v_m)_0}$
200	76	$0.7 \cdot 10^{-4}$	$0.2 \cdot 10^{-4}$	0.3
280	51	$1.7 \cdot 10^{-4}$	$10 \cdot 10^{-4}$	6.0
350	19	$1.0 \cdot 10^{-4}$	$15 \cdot 10^{-4}$	15.0

An interesting result is obtained when a single crystal of zinc covered by a film of liquid tin is maintained for a sufficiently long time without stress at 400°C, and is then placed into liquid lead at the same temperature. If the crystal is then kept in liquid lead for a sufficiently long time the effect disappears completely and the normal strength of single crystals of zinc is reestablished. This experiment indicates that no appreciable role is played by the normal diffusion of tin into zinc. Evidently the penetration of tin and the associated reduction in strength occurs only along the developing surface defects incipient shear and fracture.

We have also obtained considerable reductions in strength and plasticity in single crystals of cadmium in the presence of liquid tin (Fig. 4). The action of liquid tin on single crystals of cadmium is the same as on single crystals of zinc. Our results indicate directly the very important fact that the action of surface-active metallic fluxes is not associated with the presence of boundaries between grains of a high-melting-point metal. This action is observed also in the absence of intercrystalline boundaries in the case of metallic single crystals and may attain considerable magnitude under favorable temperature conditions. Of course, this conclusion does not mean that the grain boundaries in a polycrystalline metal play no role at all in the effect of flux action. They may play a very considerable, and in some cases a decisive role; however, we should bear it in mind that a surface-active metallic flux may reduce the strength and the plasticity of an individual grain as much as desired, owing to the sharp reduction in the interphase surface tension.

In the presence of such a strong surface-active substance as a metallic flux, even very small normal stresses in high-melting-point metals lead to the formation and evelopment of cracks that result in brittle fracture of the metal. For this it is merely necessary that the kinetics of the development of the crack correspond to the rate of penetration of the surface-active flux along these cracks into the metal undergoing deformation.

In earlier papers on the effect of ordinary organic surface-active substances on the creep of

TABLE II. Rate of steady creep at 350°C of zinc single crystals covered by alloys of tin with lead of different concentrations

PbSn alloy	$(v_m)_0 \cdot 10^4$ (min^{-1}) uncoated	$(v_m)_A \cdot 10^4$ (min^{-1}) coated	$\frac{(v_m)_1}{(v_m)_0}$
90% Pb—10% Sn	1.3	1.7	1.3
80% Pb—20% Sn	1.0	3.0	3.0
50% Pb—50% Sn	1.1	4.5	4.0
20% Pb—80% Sn	3.3	20.0	6.0

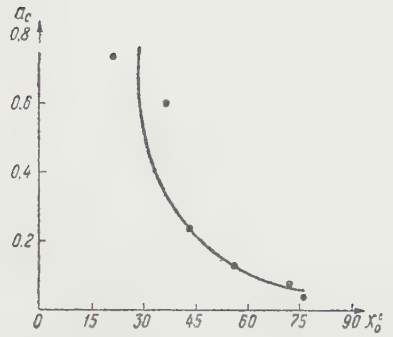


FIG. 5. Decrease in the value of the limiting plastic shear (a_s) for zinc single crystals as the initial angle of orientation of the base plane (χ_0) is increased at −196°.

metallic single crystals it was found that the rate of creep is considerably increased by adsorption of surface-active molecules contained in the surrounding medium. In connection with this it is of considerable interest to study the creep of metallic single crystals in such strongly surface-active media as fluxes of low-melting-point metals.[17]

In the presence of a tin coating, starting with 280°C, an appreciable increase in the rate of steady-state creep is observed for single crystals of zinc. Table I gives data on the change in the rate v_m of steady-state creep of single crystals of zinc at different test temperatures.

Table II gives data for the rate of steady-state creep of single crystals of zinc coated with alloys of tin with lead in different concentrations.

Tin flux, which is a very strongly surface-active for zinc, reduces sharply the level of normal fracture stresses under conditions of elongation at a constant rate of deformation, which leads to a considerable reduction of plasticity (value of deformation before rupture) and strength. However, under conditions of creep, when the constant stress acting on the metal is lower than the stress of brittle fracture, a plasticizing effect due to adsorption appears in the presence of a surface-active flux and manifests itself in a considerable increase in the rate of creep of the metal.

N. V. Pertsov has shown that mercury, like tin, behaves like a very strong surface-active substance with respect to zinc single crystals and sharply reduces their strength and particularly their plasticity.[21] Even at room temperature single crystal zinc wires become brittle under the action of mercury, i.e., they become characterized by a very small plastic elongation, which ends in brittle fracture along the basal plane (or along the first-order prismatic plane) at not very high tensile stresses.

Since molten metals reduce appreciably the strength and the plasticity of metals of higher melting points by making them brittle, it becomes necessary to make a comparative study of the laws of deformation and fracture of metals in strongly surface-active fluxes, and of similar laws for these metals in a brittle state but in nonactive media.[19]

As is well known, single-crystal zinc wires are very plastic at room temperature, i.e., they are capable of considerable plastic elongations (up to 600%). Brittle fracture along the basal plane does not occur under these conditions (at not very high rates of elongation). The brittle state may be realized in single crystals of zinc by a considerable decrease in temperature. This brittle state has the peculiarity that brittle fracture under the action of normal stresses is always preceded by a fairly considerable shear deformation along the active slip system (basal planes).

A study of the laws of brittle fracture of single crystals of zinc with different initial orientations of the basal plane with respect to the axis of the wire ($13° \leq \chi_0 \leq 80°$) was carried out at the temperature of liquid nitrogen (−196°C) under tension with a constant rate of elongation ($\cong 12\%$ min^{-1}). The critical shearing stress in the basal plane at the temperature of liquid nitrogen reaches ~130 g/mm^2 and, as has been shown by extensive experimental material,[22] does not depend on the orientation of this plane relative to the crystal axis (on the value of χ_0).

A study of the laws of brittle fracture (fracture along the basal plane) of single crystals of zinc of various orientations has led first to the discovery that the value of the plastic shear as preceding fracture increases as χ_0 decreases (Fig. 5).

TABLE III. Values of a_S, p_S, τ_S at the instant of brittle fracture along the basal plane resulting from elongation of zinc single crystals at −196°C

χ_0	a_S	p_S, g/mm²	τ_S, g/mm²
21	0.74	93	311
36	0.60	160	314
43	0.24	182	238
56	0.13	249	201
72	0.08	268	128
76	0.04	431	126

At the same time the normal fracture stresses in the basal plane (the slip plane) decrease sharply as the initial angle of orientation χ_0 of this plane is reduced or, which is the same thing, as the plastic shear preceding fracture is increased. Figure 6 shows the dependence of the normal stresses in the basal plane on the magnitude of the shear occurring in the course of deformation of single crystals of zinc with different initial orientations of the basal plane. Here it is necessary to point out that the area of the basal plane (the slip plane) remains constant to a sufficient degree of accuracy during the process of plastic shear even in the case of quite considerable deformations (up to 600%). The crescent-shaped portions of the slip planes uncovered by the shear amount to some hundredths of a percent of the area of the basal plane and, consequently, this change in area can always be neglected. The dotted curve in Fig. 6 joining the end points (corresponding to brittle fracture along the basal plane) of the individual stress-strain curves of variously oriented zinc single crystals determines, by this very fact, the dependence of the normal fracture stresses along the basal plane on the magnitude of the shear preceding the fracture. This dependence, which has been checked by us on a large number of samples, shows that the normal fracture stresses decrease sharply with increasing shear deformation (along the basal plane) that precedes the brittle fracture (along this plane).

Table III gives values of the normal fracture stresses p_S, limiting values of the magnitude of the shear a_S and values of the shear stresses τ_S as a function of χ_0.

For a proper understanding of this variation it should be pointed out that the crystallographic conditions for the deformation of zinc single crystals, together with the law that the critical shearing stress is independent of the angle of orientation of the basal plane, cause the normal stresses on the

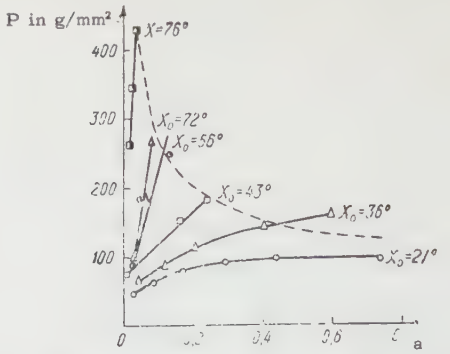


FIG. 6. Dependence of the normal stresses across the basal plane (p) on the magnitude of the shear (a) during the process of deformation of zinc single crystals with different orientations of the basal plane (χ_0) at −196°.

basal plane to be relatively small for small values of χ_0, while for large values of χ_0 these stresses increase considerably.

Since the process of plastic shear along the slip plane (the basal planes) is accompanied by a decrease in the angle χ (as a result of a deformation $\chi < \chi_0$), then for crystallographic reasons the normal stress along the basal plane must decrease in the course of the slip process. However, an increase in the shearing stress due to hardening leads to an increase in the total tensile stress, which also gives rise to an increase in the normal stress along the shear plane which exceeds the decrease in this stress for crystallographic reasons.

Figure 7 presents the dependence of the limiting plastic shear and of the normal and shearing fracture stresses on the initial angle or orientation of the basal plane. Here we can clearly see the regular decrease in normal fracture stresses corresponding to an increase in the plastic shear preceding fracture. As this deformation increases, the value of the shearing stresses also increases, and this is associated with shear hardening.

From these data it follows first of all that the so-called Zonke's law with respect to the constancy

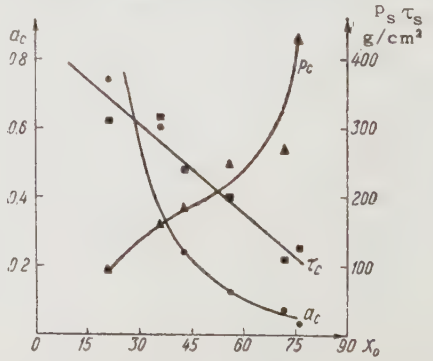


FIG. 7. Dependence of the limiting plastic shear (a_s), and the normal (p_s) and shearing (τ_s) fracture stresses on the initial angle of orientation of the basal plane (χ_0) of zinc single crystals at −196°.

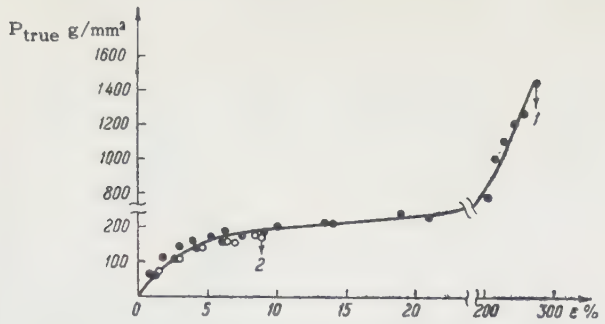


FIG. 8. Dependence of the stress (P) on the relative elongation (ε) of non-amalgamated (1) and amalgamated (2) zinc single crystals ($\chi_0 = 48°$) at 20°C.

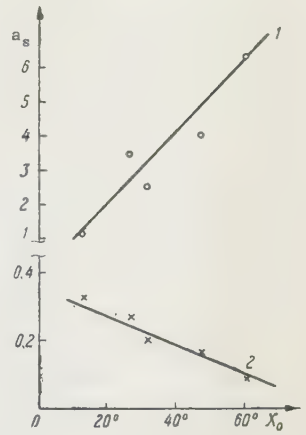


FIG. 9. Dependence of the limiting plastic shear (a_s) of non-amalgamated (1) and amalgamated (2) zinc single crystals on the initial orientation of basal plane (χ_0) at 20°C.

of normal fracture stresses does not hold for brittle single crystals of zinc. This is apparently explained by the fact that at low temperatures brittle fracture of single crystals of zinc is always preceded by an appreciable plastic shear deformation along the slip planes, while in crystals of rock salt, in which this law was discovered, plastic deformation preceding fracture is practically equal to zero. With respect to the work of E. Schmid, who confirmed the validity of Zonke's law for single crystals of zinc at low temperatures, we can say that the result obtained by him can be explained by choosing a comparatively narrow range of orientation of the basal plane in his experimental samples.[22]

At the same time it follows from the data obtained that plastic shear leads to the appearance of defects in the crystal structure which are points of incipient fracture. Even in their early stages of development these structural defects lead to brittle fracture along the slip plane (basal plane) if the level of normal stresses is sufficiently high (for large values of χ_0). However, if the magnitude of normal stresses is small (low values of χ_0), then the development of the defects up to the critical value corresponding to the given level of normal stresses occurs at a larger value of plastic shear.

The observed laws of brittle fracture of single crystals of zinc at low temperatures hold also when the transition into the brittle state is achieved not by lowering the temperature, but by the action of a strong surface-active medium (mercury), in which the deformation of crystals takes place. However, an essential difference between these brittle states is that in the surface-active medium (mercury) the brittle strength of single crystals (the magnitude of the normal fracture stresses) is sharply reduced (by a factor of 3 to 5) while the value of the shear stresses remains practically unaltered.

Mercury was deposited on the surface of the single crystals of zinc in a thin layer ($\sim 5\mu$ thick) by immersing the crystals into a saturated solution of mercury nitrate (the higher salt) and holding them in the solution for about 1 minute. Figure 8 shows typical stress-strain curves for amalgamated and non-amalgamated single crystals of zinc ($\chi_0 = 48°$) at room temperature. Here one can clearly see the sharp increase in the brittleness of crystals in the presence of mercury, but at the same time the stress-strain curve for the amalgamated sample coincides with the stress-strain curve for the non-amalgamated crystal.

Figure 9 shows the dependence of the limiting plastic shear a_S preceding fracture of non-amalgamated and amalgamaged crystals on the initial orientation of the basal plane. While in the absence of mercury this limiting shear increases with increasing χ_0, (i.e., for a more nearly transverse orientation of the basal plane the crystals turn out to be more plastic), the amalgamated samples behave in a completely opposite manner. Such a radical change in the nature of the functional dependence $a_S = a_S(\chi_0)$ under the action of mercury is associated with the transition of the zinc single crystals in the presence of mercury from the plastic to the brittle state, and with the resultant new laws governing deformation and fracture.

Indeed, the dependence shown in Fig. 10 of the

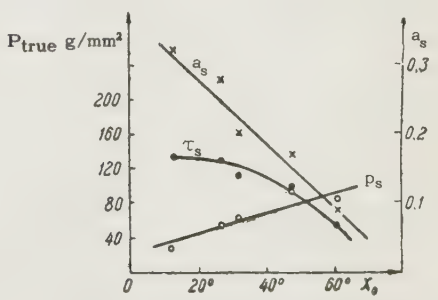


FIG. 10. Dependence of the limiting value of the shear (a_s), and the normal (p_s) and shearing (τ_s) fracture stresses on the initial orientation of the basal plane (χ_0) of amalgamated zinc crystals at 20°C.

TABLE IV. Values of a_S, p_S, τ_S at the instant of brittle fracture along the basal plane of amalgamated zinc crystals at 20°C

χ_0	a_S	p_S, g/mm²	τ_S, g/mm²
13	0.33	29	136
27	0.28	58	131
32	0.20	64	114
48	0.17	90	96
61	0.09	85	55

limiting value of the shear, and of the normal and shearing stresses at the moment of fracture, on the orientation of the basal plane is completely analogous to a similar dependence for single crystals of zinc at low temperatures (Fig. 7). Table IV lists values of the limiting shear deformation, and of the shear and normal fracture stresses along the basal plane for brittle amalgamated samples of different orientations at room temperature.

While the normal fracture stresses is lower (by a factor 3 to 5) than in non-amalgamated crystals (at −196°C), the shearing stresses in the basal plane of amalgamated samples remain the same as in the case of non-amalgamated single crystals of zinc at the same room temperature. However, at the temperature of liquid nitrogen the amalgamated samples turned out to be no less plastic than the non-amalgamated ones, while the normal and shearing stresses at the instant of brittle fracture along the basal plane turned out to be even higher in the presence of the mercury film than in its absence (Fig. 11). Since the amalgamated samples were immersed into liquid nitrogen 2 or 3 minutes after the mercury film had been applied and, consequently, there was no time for any appreciable diffusion penetration of mercury into zinc, the effect of the increase of strength of amalgamated samples at low temperatures can be explained by the appearance of a thin polycrystalline coating of mercury on the zinc single crystals, which hinders plastic shear and brittle fracture along the basal plane. Hardening of amalgamated samples in liquid nitrogen can also be observed as a result of the diffusion penetration of mercury into zinc (alloying) if the amalgamated samples are maintained at room temperature for a sufficiently long time prior to the experiment (~48 hours).[23]

Thus, mercury used as a surface-active substance produces in zinc single crystals a brittle state characterized by the same general laws of deformation and fracture as the brittle state due to a lowering of temperature in the absence of surface-active substances. The sharp reduction

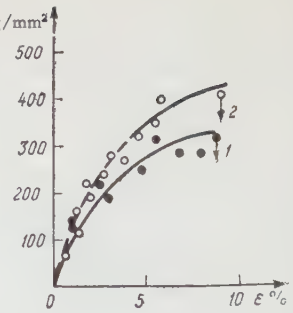


FIG. 11. Dependence of the stress (P) on the relative elongation (ε) of non-amalgamated (1) and amalgamated (2) zinc single crystals ($\chi_0 = 48°$) at −196°C.

in the normal fracture stresses in the presence of mercury is due to an appreciable reduction in the interphase surface tension at the boundary between the zinc and the saturated solution of zinc in mercury. The structure defects (micro-cracks) formed during deformation are rapidly filled with mercury by two-dimensional migration, and this greatly facilitates their further development until brittle fracture can occur under the action of even small normal stresses. The high mobility of mercury atoms, which enables them to penetrate sufficiently rapidly into the forming micro-cracks, is a necessary condition for the reduction in strength and plasticity. At low temperatures, when the mercury atoms are not mobile, one observes an increase in the brittle strength of zinc single crystals covered by a mercury film. This is due either to alloying of the zinc with the mercury, or to the resistance of the film itself to shear and fracture.

The experiments described above,[18,19,21] as well as numerous data of other authors, indicate in a convincing manner that brittle fracture of a metallic single crystal is always preceded by plastic deformation of some kind; the stress p_S normal to the cleavage plane at the moment of fracture is by no means constant, but depends strongly on the magnitude of the preceding plastic shear a_S, and at the same time on the shearing stress τ_S in the slip plane attained at the moment of fracture. In our case of brittle fracture of zinc single crystals when the basal plane is simultaneously the slip plane and the cleavage plane, the interdependence between p_S and τ_S is particularly clearly evident: as the value of p_S increases crystal fracture occurs at smaller values of τ_S (Tables III and IV). In this case one should apparently speak not of the deviations from Zonke's law which have been repeatedly mentioned earlier, but of a complete inapplicability of this law.

It is well known that the actual fracture stress p_S along the cleavage planes is several orders of magnitude lower than the so called "theoretical" value $p_{theor} \cong \sqrt{E\sigma/b}$, calculated on the basis of some model of interatomic forces for an ideal

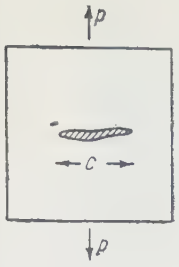
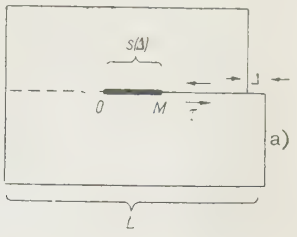
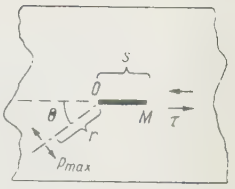


FIG. 12. Illustration for the calculation of p_s (c).





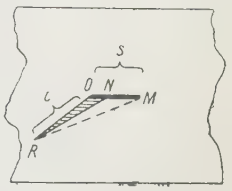


FIG. 13. (a) Appearance of an incomplete local shear in a glide plane. (b) Concentration of normal stresses in the neighborhood of the region of an incomplete local shear. (c) Appearance of a micro-crack in the neighborhood of maximum concentration of normal stresses.

crystalline lattice containing no defects;[25,26] here E is Young's modulus, σ the specific surface free energy, and b the lattice constant. Such a discrepancy is usually associated with the presence of various structural defects, primarily micro-cracks, in the actual crystal. The inapplicability of Zonke's law, and the observed dependence of p_s on τ_s and a_s in the case of brittle fracture of single crystals of zinc, enable us to take the following step in the explanation of the mechanism of brittle fracture.

It is natural to suppose that if the fracture is due to the presence of a micro-crack the magnitude c of which becomes "dangerous" for a given p_s, the initiation and the growth of the micro-crack up to the dangerous value are associated with the preceding plastic deformation, in the course of which inhomogeneities of shear lead to sharp local stress concentrations and to the appearance of micro-cavities.[6,24]

It is therefore necessary first to find the value of the critical normal stress p_c which is dangerous in the presence of a crack of size c and then to try to estimate the maximum size c of the crack that may appear in the crystal at a given shearing stress τ.[24]

Following Zener,[27] the simplest and most obvious estimate of the dangerous tensile stress p_c in the presence of a crack of size c can be obtained in the following manner (Fig. 12). Prior to the appearance of the crack the density of elastic energy in the solid is $w = p^2/2E$. When the crack appears the stress is removed from an area approximately equal to c^2; the decrease in elastic energy (per unit length in the direction perpendicular to the plane of the diagram) amounts in this case to $\cong p^2c^2/2E$. At the same time the formation of the new surface of approximate size 2c is associated with an expenditure of work equal to $2c\sigma$. The total increment in energy is equal to $\Delta W = 2c\sigma - p^2c^2/2E$. The curve $\Delta W(c)$ has a maximum at $(\partial/\partial c)\Delta W = 0$. Consequently, a crack of width c becomes unstable when a "dangerous" stress $p_c = \sqrt{2\sigma E/c}$ is applied perpendicular to its plane.

A similar formula was first proposed by Griffith.[28] Subsequently several authors have shown that, with the exception of some differences in the numerical coefficient, the result remains the same under very varied formulations of the problem: both for thin and for thick slabs (two-dimensional problem), in the case of a three-dimensional problem (crack in the form of an ellipsoid),[29] and also in going over from the calculation of the energy to a direct determination of stress concentrations at the edge of the crack,[28,30] in particular in the analysis of interatomic forces and of the location of atoms.[31]

Thus, one can expect that, in the presence of a crack of size c in the crystal, the dangerous stress normal to the plane of the crack is given by the relation

$$p_s = \alpha \sqrt{\frac{E\sigma}{c}}, \qquad (1)$$

where α is a certain dimensionless coefficient not much different from 1 in order of magnitude. (However, in going over from isotropic media, for which Eq. (1) is derived, to crystals with very pronounced cleavage, the values of α are apparently appreciably less than unity).

Let us assume that under the action of a shearing stress τ the crystal has undergone a shear in the slip plane by an amount Δ which, however, did not spread over the shole cross section of the crystal, owing to the presence of some sufficiently strong obstacle O in the slip plane (Fig. 13). In this case the main region of localization of the shear turns out to be the interval MO = s along the slip plane (the value of s obviously depends on the initial shear Δ). The stress field around such a defect can be compared to the elastic field

of a cut of length s parallel to the applied shearing stress τ. As is shown in the theory of elasticity (cf., for example, Neiber[32]) the maximum tensile stresses occur in this case along a line that makes an angle Θ with the plane of the cut, where they attain a value $p_{max} \cong \sqrt{s/r}$ (r is the distance between the given point and the vertex of the cut).

We assume that a wedge-shaped crack NOR of magnitude $NR \cong OR = c$ originates in the region of maximum stress concentration (the base of the wedge is approximately equal to the initial shear Δ). Then the high stress concentration will be relieved not only along the surface of the crack NOR itself, but also over a considerably larger area, $\cong \Delta MOR$.

In the absence of the crack the density of elastic energy in the region of the interval MO is $w = p^2/2E \cong \tau^2 s/2Er = w(r)$. The decrease in energy as the crack opens up (per unit length in the direction perpendicular to the plane of the diagram) is approximately given by

$$\iint_{\Delta MOR} w(r)\,dS \cong \int_{r_0}^{c} \frac{\tau^2 s^2}{2Er}\left(1-\frac{r}{c}\right)dr$$

$$= \frac{\tau^2 s^2}{2E}\left(\ln\frac{c}{r_0} - \frac{c-r_0}{c}\right) \cong \frac{\tau^2 s^2}{2E}\ln\frac{c}{r_0}$$

where $c \gg r$. (The lower limit of integration $r_0 \cong b$ is determined by the inapplicability of the equations of the theory of elasticity in the immediate neighborhood of the point O.)

At the same time the formation of the crack requires an expenditure of work $\cong 2c\sigma$. The total increase in energy is equal to $\Delta W = 2c\sigma - (\tau^2 s^2/2E)\ln(c/r_0)$. The function $\Delta W(c)$ has a minimum at $(\partial/\partial c)\Delta W(c) = 2\sigma - (\tau^2 s^2/2E)(1/c) = 0$. Consequently, as incomplete shears are accumulated in the slip plane in the interval $s = s(\Delta)$, a wedge-shaped equilibrium crack of width $c \cong \tau^2 s^2/4E\sigma$ may appear in the crystal.*

From the point of view of an analysis of brittle fracture, we are interested in the limiting case of the largest possible c. We have such a case if in front of the "completely insurmountable" obstacle O the interval s of high concentration of incompleted shears reaches the same order of magnitude as the dimensions of the whole slip plane L (in the case of not very thick single crystals, ≤ 1 mm, L is approximately equal to the diameter; in the case of polycrystalline samples, L equals the grain size).

*For such an equilibrium crack actually to appear a gain of energy is necessary: $\Delta W < 0$ with $c = \tau^2 s^2/4E\sigma$, or $\tau^2 s^2/4E\sigma > 2.7r_0$, i.e., of the order of several b.

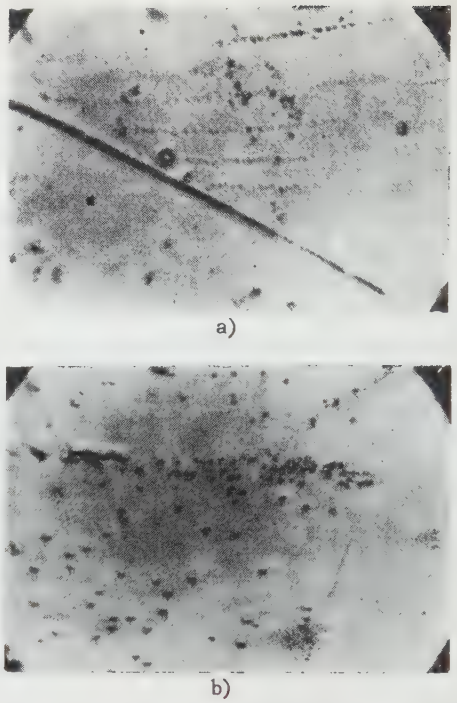


FIG. 14. Microphotographs of etch pit patterns on the surface of zinc crystals (a) and of cadmium single crystals (b); pile-up of dislocations in front of a grain boundary may be seen;[6] 800×.

Thus it would appear that when a shearing stress τ is applied in the slip plane, a crack can appear in the crystal of dimensions up to

$$c = \beta\frac{\tau^2 L^2}{E\sigma}, \qquad (2)$$

where the dimensionless coefficient β does not differ appreciably from unity in order of magnitude.* A similar relationship can be obtained in a more rigorous manner from the dislocation theory.

The process, discussed above, of the concentration of incomplete shears amounts, evidently, to none other than the formation of pile-up of dislocations stopped by an obstacle in the slip plane. According to Mott's theory such pile-ups act as the principal concentrators of the internal stresses that arise in the crystal in the course of plastic deformation.[33-35] Pile-ups of dislocations have been found experimentally, for example, in investigations of the etching of dislocations on the crystal surface (Fig. 14; cf. also Meleka[36]).

The problem of the location of n parallel edge

*From this point of view the increased strength of single crystals as a result of alloying may be due to the decrease in the effective value of s, and correspondingly of the coefficient β in (2): alloyed single crystals deform in a more homogeneous manner, and under the same stresses the incomplete local shears turn out to be smaller than in pure single crystals.

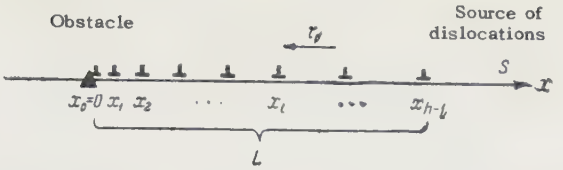


FIG. 15. Pile-up of dislocations in front of an obstacle in the slip plane.

dislocations within a pile-up has been solved for an isotropic medium by the methods of the theory of elasticity.[37] The solution is based on the well-known expression

$$F = b \cdot \tau \tag{3}$$

for the force acting in the slip plane per unit length of the dislocation line[38] (τ is the shearing stress in the direction of the Burgers vector of the dislocation b), and on the expression for the field of the elastic stresses of an edge dislocation

$$p_{rr} = p_{\vartheta\vartheta} = -\frac{Gb}{2\pi(1-\mu)} \cdot \frac{\sin\vartheta}{r}; \qquad \tau_{r\vartheta} = \frac{Gb}{2\pi(1-\mu)} \cdot \frac{\cos\vartheta}{r} \tag{4}$$

(in cylindrical coordinates the angle ϑ is measured from the slip plane; μ is Poisson's ratio; the functional form of this dependence is well confirmed experimentally[39,40]). The general outline of the argument is approximately as follows.

It is assumed that each point of the glide plane OS (Fig. 15; O is the place at which the leading dislocation has stopped in front of the obstacle, and S is the source of these n dislocations) is subject to a tangential stress $\tau(x)$ made up of an external stress in the given plane τ_0, of the field of the leading dislocation (whose position $x_0 = 0$ is fixed), and of the fields of the remaining dislocations (whose coordinates x_1 to x_{n-1} are to be determined). In accordance with (3) and (4) this stress is equal to

$$\tau(x) = \tau_0 - \frac{Gb}{2\pi(1-\mu)} \cdot \frac{1}{x} - \sum_{i=1}^{n-1} \frac{Gb}{2\pi(1-\mu)} \cdot \frac{1}{x - x_i}.$$

The conditions of equilibrium for the dislocations are given by the $n-1$ equations

$$\tau(x_j) = 0; \qquad j = 1, 2, \ldots, n-1,$$

whose solution is reduced to the solution of the algebraic equation

$$\frac{d}{dz}\left[e^z \frac{d^n}{dz^n}(e^{-z} z^n) \right] = 0,$$

where $z = x \cdot 2\tau_0 \cdot 2\pi(1-\mu)/Gb$, i.e., the desired coordinates of the dislocations are determined by the $n-1$ roots of the first derivative of the n-th order Laguerre polynomial; for large n and small z the last equation can be reduced to the Bessel

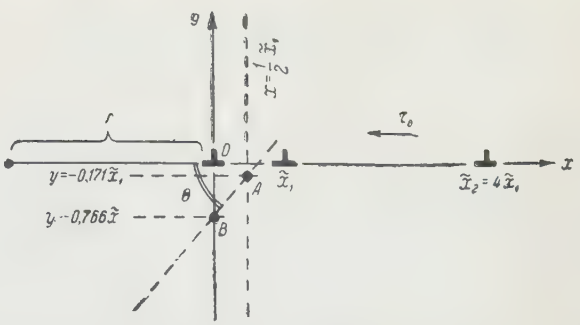


FIG. 16. Concentration of stresses in the neighborhood of the head of a pile-up of dislocations.

equation. The distance x_1 between the two leading dislocations is

$$x_1 = \frac{(3.83)^2 \cdot Gb}{4n \cdot 2\tau_0 \cdot 2\pi(1-\mu)} = 0.42 \frac{Gb}{n\tau_0}. \tag{5}$$

If the average values of i are greater than 1, but considerably less than n, the following approximation holds:

$$x_i \approx \frac{\pi Gb}{16 n \tau_0 (1-\mu)} i^2, \tag{6}$$

while the total distance occupied by the concentration of dislocations is given by

$$x_{n-1} = L \approx \frac{nGb}{\pi\tau_0(1-\mu)}. \tag{7}$$

The approximate estimate (7) is valid only for large n, $\sim 10^3$, but is too inaccurate for small n, ≤ 10.

A different way of treating the problem of a pile-up of dislocations has been indicated by Leibfried. The pile-up is approximated by a continuum of dislocations with a distributed of Burgers vector; this leads to a solution of integral equations.

A rigorous complete description of the elastic stress field of a pile-up of dislocations requires the summation of series with terms of type (4). Approximate results relating to several special cases were obtained by Koehler,[38,41] who approximated all the values of x_i, including x_1, by expression (6). The value of x_1, obtained from (6) and denoted in Fig. 16 by $\tilde{x}_1$, is equal to

$$\tilde{x}_1 = \frac{\pi Gb}{16 n \tau_0 (1-\mu)} \approx 0.28 \frac{Gb}{n\tau_0} = \frac{2}{3} x_1.$$

According to Koehler's calculations, the maximum tensile stress p_{xx} along the Oy axis occurs at point A when $y = -0.766\,\tilde{x}_1$ and is equal to $\sim n\tau_0$. Along the straight line $x = \tilde{x}_1/2$ this stress reaches a maximum at point B when $y = -0.171\,\tilde{x}_1$ and has a value of $\sim 7n\tau_0$. The author emphasizes that the large stress concentrations obtained are very important in plastic and fatigue fracture of

materials. However, Koehler considers that his results are too high by approximately 50%, owing to the known inaccuracy of the approximation for x_1. We note that the angle Θ between the direction of maximum p_{xx} and the negative Ox axis amounts to somewhat more than $\pi/4$.

It can be easily shown that similarly high concentrations in the neighborhood of a pile-up of dislocations are also attained by the shearing stresses, whose maximum values evidently lie along the extension of the Ox axis towards $\theta = \pi$. With the same simplifying assumption with respect to x_1, the tangential stress at a distance r in front of the head of the pile-up is

$$\tau(r) = \tau_0 + \frac{Gb}{2\pi(1-\mu)} \sum_{i=0}^{n-1} \frac{1}{r + i^2 \tilde{x}_1}.$$

When $r \ll \tilde{x}_1$ the sum of all the terms of the series, with the exception of $i = 0$, is given by

$$\sum_{i=1}^{n-1} \frac{1}{r + i^2 \tilde{x}_1} \approx \frac{1}{\tilde{x}_1} \sum_{i=1}^{n-1} \frac{1}{i^2} < \frac{1}{\tilde{x}_1} \cdot \frac{\pi^2}{6} \ll \frac{1}{r},$$

i.e., in the immediate vicinity of the head of the pile-up the field coincides with the field of the leading dislocation. When $r \gg L$

$$\sum_{i=0}^{n-1} \frac{1}{r + i^2 \tilde{x}_1} \approx \int_0^n \frac{d\xi}{r + \tilde{x}_1 \xi^2} = \frac{1}{\sqrt{r\tilde{x}_1}} \tan^{-1} \sqrt{\frac{n^2 \tilde{x}_1}{r}} \approx \frac{n}{r},$$

since $n^2 \tilde{x}_1 = (\pi^2/16)\, L < L \ll r$. This means that at very large distances the field of the pile-up is similar to the field of a single "large" dislocation whose Burgers vector is equal to nb.

The high stress concentration is particularly significant in the intermediate region $\tilde{x}_1 \ll r \ll L$. In virtue of the left hand side of this inequality, the summation can, as before, be replaced by an integration, while in virtue of the right hand side we have

$$\tan^{-1} \sqrt{\frac{n^2 \tilde{x}_1}{r}} \cong \frac{\pi}{2}, \quad \text{and} \quad \tau(r) = \tau_0 + \frac{Gb}{2\pi(1-\mu)} \cdot \frac{1}{\sqrt{r\tilde{x}_1}} \cdot \frac{\pi}{2}.$$

By substituting into this the value of $\tilde{x}_1$ from (8), and by utilizing (7), we obtain

$$\tau(r) = \tau_0 \left\{ 1 + \sqrt{\frac{L}{r}} \right\} \quad \text{for} \quad \tilde{x}_1 \ll r \ll L,$$

or, approximately,

$$\tau(r) \cong \tau_0 \sqrt{\frac{L}{r}}. \tag{9}$$

An approximate solution for all the components of the stress field of the pile-up in the region $x_1 \ll r \ll L$, which is of greatest interest, has been given by Stroh.[42] The distance between the two leading dislocations obtained by him and given by

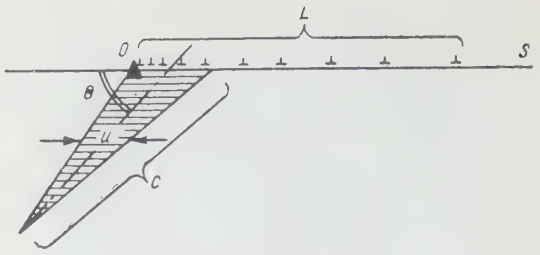


FIG. 17. Appearance of a micro-crack in the neighborhood of the head of a pile-up of dislocations.

$$\frac{Gb}{n\pi\tau_0(1-\mu)} \approx 0.46 \frac{Gb}{n\tau_0}$$

differs only little from the value of x_1 obtained from (5), while the length of the pile-up coincides exactly with the value of L from (7). For the stress components Stroh obtains

$$\left.\begin{aligned} \frac{1}{2}(p_{xx} + p_{yy}) &= \tau_0 \sqrt{\frac{L}{r}} \cdot 2 \sin\frac{\vartheta}{2}, \\ \frac{1}{2}(p_{xx} - p_{yy}) &= \tau_0 \sqrt{\frac{L}{r}} \cdot \left(2 \sin\frac{\vartheta}{2} + \sin\vartheta \cos\frac{3}{2}\vartheta \right), \\ \tau_{xy} &= -\tau_0 \sqrt{\frac{L}{r}} \cdot \frac{1}{2} \left(2 \cos\frac{\vartheta}{2} - \sin\vartheta \sin\frac{3}{2}\vartheta \right). \end{aligned}\right\} \tag{10}$$

The last of these formulas agrees, when $\vartheta = \pi$, with the approximate estimate of the largest of the shearing stresses (9). The maximum normal stresses occur, according to (10), along the straight line $\Theta \cong 70°$ and have the value

$$p_{\max} = \frac{2}{\sqrt{3}} \tau_0 \sqrt{\frac{L}{r}}. \tag{11}$$

Mott has advanced the hypothesis that the high stress concentration near the head of a pile-up of dislocations may give rise to the appearance and development of cracks.[33,43*] This idea has been developed in a number of papers by Stroh.[44-46]

By assuming that the crack is formed by the merging of all n dislocations, i.e., that it contains a dislocation with a Burgers vector nb, Stroh finds all the components of the stress field of a crack making an arbitrary angle Θ with the vector nb, and also the outline of the crack for the case $\Theta = \pi/2$: its "thickness" is given by

$$u = \frac{1}{2} nb \left\{ 1 + \frac{2}{\pi} \arcsin \frac{c-2r}{c} \right\},$$

i.e., the crack has the form of a "wedge" with a base nb, which lies in the glide plane, a result

*Other dislocation models have also been proposed for the origin of cracks. Thus, Fujita[47] bases his discussion on the consideration of two opposing accumulations of dislocations of opposite signs in neighboring slip planes, while Fisher[48,49] bases his discussion on an analysis of the pile-up of vacancies arising as a result of intersection of dislocations containing screw components.[50]

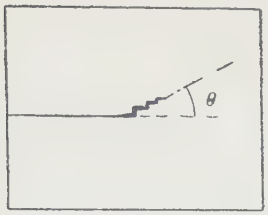


FIG. 18. A schematic representation of the alternation of cleavages along different planes.

which is in excellent agreement with the concept of an edge dislocation interpreted as an extra half-plane inserted at right angles to the glide plane.

By utilizing the method of making an imaginary cut in the solid and pushing apart the edges of the cut, and by utilizing the found components of the field of the crack, Stroh determines its energy (per unit length at right angles to the plane xOy):

$$W = \frac{n^2Gb^2}{4\pi(1-\mu)} \ln\frac{4R}{c} + 2c\sigma,$$

where R is the distance from the crack to the surface of the solid. The author emphasizes that W does not depend on Θ. This expression has a minimum given by

$$W_{cr} = \frac{n^2Gb^2}{4\pi(1-\mu)} \ln\frac{32\pi e(1-\mu)R\sigma}{n^2Gb^2},$$

which corresponds to an equilibrium width of the crack

$$c = \frac{n^2Gb^2}{8\pi(1-\mu)\sigma}. \qquad (12)$$

Stroh[46] makes use of this result for a quantitative explanation of the experimental data[51] on the increase in the electrical resistance and on the decrease in the density of nickel after extensive cold working.

Utilizing (7) to estimate the maximum number of dislocations in front of the obstacle we obtain

$$c \approx \frac{\pi(1-\mu)}{8} \cdot \frac{L^2\tau_0^2}{G\sigma},$$

which completely agrees with expression (2).

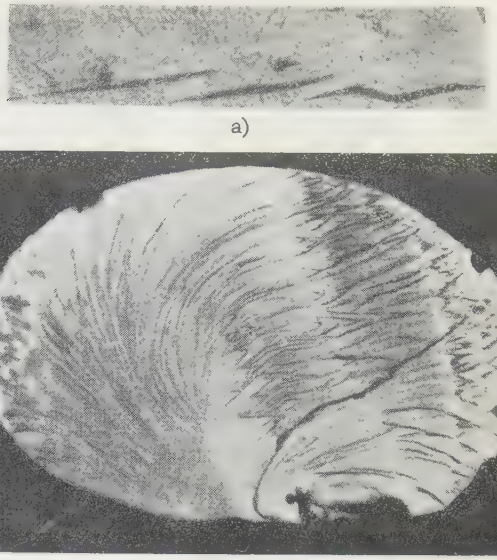
Now on substituting expression (2) obtained for the maximum possible size of the crack into Griffith's relation (1) we obtain in place of Zonke's law p_S = const the new relation

$$p_S\tau_S = \text{const} = K^2, \qquad (14)$$

where

$$K = \gamma' \sqrt{\frac{E\sigma}{L}} = \gamma \sqrt{\frac{G\sigma}{L}}; \qquad (15)$$

here G is the shear modulus, while the dimensionless coefficients γ' and γ are close to unity in order of magnitude. Relation (14) may be characterized as "the condition for the constancy of the product of normal and shearing stresses" in the case of brittle fracture.[24]



a)

b)

FIG. 19. (a) Microphotograph of wavy cracks in a zinc crystal; 200 ×. (b) Microphotograph of steps[6] arising on the surface of a zinc single crystal undergoing brittle fracture; 50 ×.

It should be emphasized immediately that the stresses p_S and τ_S in relation (14) are applied, generally speaking, to different planes (cf. Fig. 13b). However, in the case of zinc single crystals conditions are considerably simplified, since the only slip plane and the only pronounced cleavage plane (under given conditions) both coincide with the same basal plane.

It is advantageous for the crack to develop in that direction Θ where the local concentration of tensile stresses is the highest (according to Stroh[42] Θ ≅ 70° in a one-dimensional isotropic medium). Since in a hexagonal zinc crystal there is no cleavage plane in the appropriate direction, the initial micro-crack must be like a staircase in shape: cleavages along the basal plane alternate most probably with cleavages along a prismatic plane (Fig. 18). (Such staircase cracks can be seen, for example, in Fig. 19a; Fig. 19b may serve as an example of repeatedly observed[6] small steps on the mirror-like surface of a brittle shear.) But it is much easier for a cleavage to occur along a basal plane rather than along a prismatic plane, and from this point of view it is more advantageous for the "basal" steps to be appreciably longer than the "prismatic" steps. Thus, one can expect that in the case of zinc the angle Θ will not be large even during the initial stage of development of the crack, while Θ becomes equal to zero after Griffith's criterion (1) has been attained. Consequently, in the analysis of data on brittle fracture of zinc single crystals along the basal plane the stress p_S

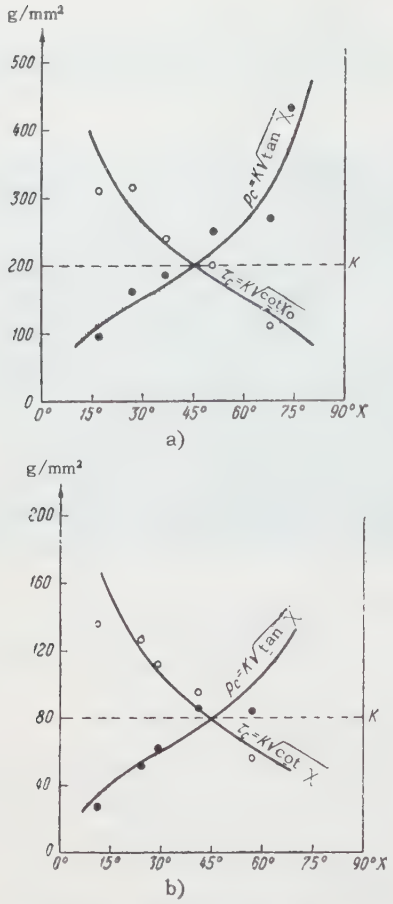


FIG. 20. Normal (p_s) and shearing (τ_s) stresses accompanying brittle fracture of zinc single crystals: (a) non-amalgamated single crystals at −196° C; (b) amalgamated single crystals at room temperature. The curves have been constructed on the basis of theory.

under the condition (14) should be referred to the same plane as τ_S.

If fracture occurs at an angle χ between the slip plane and the direction of tension then $\tau_S = P_S \sin\chi\cos\chi$ and $p_S = P_S \sin^2\chi$, where P_S is the true tensile stress at the instant of fracture. Consequently, for purely geometrical reasons, we have

$$\frac{p_s}{\tau_s} = \tan\chi. \tag{16}$$

On solving (16) simultaneously with (14) we obtain

$$\left.\begin{aligned} p_s &= K\sqrt{\tan\chi}, \\ \tau_s &= K\sqrt{\cot\chi}. \end{aligned}\right\} \tag{17}$$

Figure 20a shows that the experimental values of $p_S(\chi)$ and $\tau_S(\chi)$ in the case of brittle fracture of zinc single crystals along the basal plane at the temperature of liquid nitrogen agree well with curves corresponding to equations (17), with $K = 200$ g/mm^2. On assuming for zinc samples $G = 3 \times 10^{11}$ dynes/cm^2, $\sigma = 10^3$ ergs/cm^2, and the diameter $L = 0.1$ cm, we obtain $\gamma = K\sqrt{G\sigma/L} = 0.37$.

Figure 20b presents analogous data for brittle fracture of amalgamated zinc single crystals at room temperature. The experimental points, as before, agree within experimental error with the curves (17), but the constant K turns out to be approximately 80 g/mm^2, which apparently can be associated only with a sharp decrease in the surface free energy of zinc in the presence of mercury. This relative change in σ amounts to $(K_{Zn}/K_{Zn\text{-}Hg})^2 \cong 6$ fold (more accurately, somewhat less, if we take into account the dependence of the modulus G on the temperature), i.e., in the presence of mercury the surface tension in the incipient micro-cracks amounts to approximately 150 to 200 ergs/cm^2.

A convincing experimental confirmation of the linear dependence of the brittle-fracture stress on $\sqrt{\sigma}$ is given by experiments[52] on the study of the hydrogen brittleness of steel.

Orowan[53] ascribes the decrease in the strength of glass with increasing time of application of the load to a reduction in the surface tension due to the adsorption of gases from the atmosphere. Indirect experimental data enable us to assume that in this case also the strength is reduced in proportion to the square root of σ.[54]

Condition (14) can be written in the following way: $p_S\tau_S = P_S^2 \sin^3\chi\cos\chi = K_2$, where P_S is the tensile stress at fracture. At $\chi_{opt} = 60°$ P_S has a minimum, $P_{c\,min} = 1.76\,K$. It is natural to suppose that in the case of brittle fracture of a polycrystalline sample the first dangerous crack occurs in the grain for which χ is closest to χ_{opt}; the condition for brittle fracture of a polycrystalline sample is then

$$P_s = 1.76\,K = \gamma_p\sqrt{\frac{G\sigma}{L}}, \tag{18}$$

where the dimensionless coefficient γ_p does not differ appreciably from unity in order of magnitude; here L should be interpreted as the grain size. One should expect γ_p to be larger than the corresponding factor γ in the case of a single crystal, not only because of the "geometrical" factor 1.76, but also as a result of the influence of the neighboring grains (a polycrystalline sample with grain L is, generally speaking, stronger than a single crystal of diameter L). Stroh[45] and Petch[54] arrive at a similar dependence of P_S on L from the point of view of dislocation theory.

Petch[54] starts from the following simple considerations. The number of dislocations in the pile-up $n \sim$ (is proportional to) $L\tau_0$ [cf. (7)]; the maximum tensile stress in the region of the vertex of the pile-up[38] is $p_{max} \sim n\tau_0 \sim L\tau_0^2$ or, more accu-

rately, $\sim L(\tau_0 - \tau_i)^2$, where τ_i denotes the internal opposing field that prevents the motion of dislocations ("friction" due to the interaction between dislocations and point structural defects). Fracture occurs when p_{max} reaches the "theoretical" value p_{theor}. From this we obtain $p_{theor} \sim L(\tau_S - \tau_i)^2 \sim L(P_S - P_i)^2$, i.e., $P_S = P_i + \text{const}/\sqrt{L}$. Experiment shows[54] that in the case of zinc the term P_i is very small, and $P_S \cong \text{const}/\sqrt{L}$.

Stroh[45] compares the results obtained by him for the minimum energy of the field of the crack, $W_{cr} = \frac{n^2Gb^2}{4\pi(1-\mu)} \ln \frac{32\pi e(1-\mu)R\sigma}{n^2Gb^2}$, and for the pile-up of dislocations prior to the formation of the crack

$$W_{gr.\,d.} = \frac{n^2Gb^2}{4\pi(1-\mu)} \ln \frac{4\pi\sqrt{e}(1-\mu)\tau_0 R}{nGb}.$$

The crack originates when $\Delta W = W_{cr} - W_{gr.d.} < 0$, i.e., when

$$n\tau_0 b \geqslant 8\sqrt{e}\,\sigma. \qquad (19)$$

Using (7) to eliminate n and setting $\tau_S \sim P_S$, we obtain $P = \text{const}/\sqrt{L}$, where const $\cong 10^8$ dynes/cm$^{3/2}$.

Experimental data[45] confirm the characteristic dependence $P_S \sim 1/\sqrt{L}$, with the constant having a value 2.3×10^8 dynes/cm$^{3/2}$ in the case of iron and 1.0×10^8 dynes/cm$^{3/2}$ in the case of zinc. The latter means that $\gamma_p \approx 6$ in expression (18).

Stroh's criterion (19), which determines for a given τ_0 the magnitude of the "dangerous" pile-up of dislocations n, is of particular interest. In connection with this it is necessary to point out certain contradictions in Stroh's scheme. His criterion has been obtained essentially as the condition under which it becomes advantageous for all n dislocations to merge and to form an equilibrium crack of dimensions $c \cong n^2b$. The calculations are awkward and cannot be regarded as rigorously founded, while the scheme itself excludes consideration of a gradual development of the crack and requires sudden formation of cracks of appreciable size c. In spite of the fact that condition (19) was obtained for an equilibrium crack, Stroh uses it as the condition for fracture; in doing this he pays attention only to the shearing stress and does not take into account the normal component of the applied stress, i.e., the orientation χ. However, comparison with experiment[45] shows that condition (19) holds for polycrystalline samples (when the dependence on χ disappears), and, moreover, it is specifically the condition that the crack should no longer be an equilibrium one.

Such a conclusion is reached also in the scheme proposed by us: by eliminating L from equation (17) with the aid of (7) and by denoting the "dangerous" value of n by N we obtain

$$N\tau_s b = k\sigma \cot\chi, \qquad (20)$$

where the dimensionless coefficient k is close to unity in order of magnitude. Now the above criterion already has a definite sibnificance as a condition for the fracture of the crystal, i.e., of the attainment of Griffith's criterion by a crack which was previously an equilibrium crack. At the same time, since we have considered the roles played by the shearing and by the normal stresses separately, we have introduced into equation (19) a dependence on the orientation of the single crystal.

The earlier stages of development of a microcrack can be represented in the following manner.[6] Already at small shearing stresses $\lesssim 10^7$ dynes/cm^2 the pile-ups of dislocations in different slip planes of the single crystal (in front of sufficiently strong obstacles) may attain values of $n \sim 10^2$ to 10^3. On making use of relations (3) to (5) and (8), the distance between the leading dislocations is found to be reduced to a few multiples of b, becoming smaller than the width of the dislocation,[55] while the forces of repulsion between them exceed the so-called "theoretical" value of shearing stress $\cong G/2\pi$ to $G/30$.[56,61] This means that the linear theory no longer holds in the immediate vicinity of the head of the pile-up. It turns out to be advantageous for the frontal dislocations to merge and to form a hollow dislocation nucleus in accordance with the scheme:

The potential barrier to be overcome is particularly small because the initial hollow nucleus does not yet have a developed surface, i.e., σ is still very small. As τ_0 increases, an increasingly large number of dislocations n may accumulate in front of the obstacle, and an increasingly larger fraction n_S may enter the developing hollow nucleus. As the nucleus attains a size $\cong$ 5b to 10b, a crack begins to develop gradually sideways from the nucleus at some average angle Θ to the slip plane. The scheme that has led us to relation (2) now becomes valid. Even at low stresses, hollow dislocation nuclei and incipient micro-cracks are capable of producing rapid irregular diffusion — a two dimensional migration of the active adsorbable substance into the crystal; indeed, a single crystal of zinc exhibits great brittleness and weakness im-

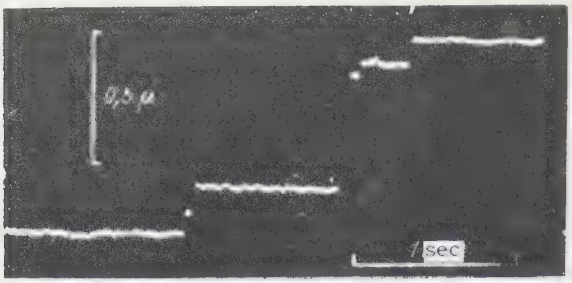


FIG. 21. Oscillogram of staircase deformation of a cadmium single crystal.[6]

mediately after being coated with a mercury film.

Investigations of elementary discontinuous deformations of different metallic single crystals elongated under a constant load near the plastic flow limit[6,57,58] confirm the existence of small equilibrium cracks, which are far from the "dangerous" size, in plastic single crystals undergoing deformation. To carry out these experiments special apparatus was constructed [59,6] to record simultaneously small deformation discontinuities (with an accuracy down to 50 or 100 A) and the accompanying increments in electrical resistance of the samples (with an accuracy down to several thousandths of a microhm); the moving parts of the apparatus have a very low intertia. The main results of such experiments[6,57,58] can be summarized in the following way:

1. The maximum value of an elementary deformation discontinuity, i.e., one which is localized in one slip zone, is 700 to 1000 A in the case of single crystals of zinc, cadmium, and tin 0.5 to 1.0 mm in diameter and for average values of the angle χ (we call such steps "single" steps). By regarding such a step as an avalanche release of pile-up of dislocations we find that the maximum number of dislocations in the pile-up amounts under the given conditions to $n \cong 10^3$. This agrees completely with the results of the study of slip lines[60] and enables us to obtain a more precise value for the numerical coefficient in (7).

2. The increment in the electrical resistance due to a single deformation step is smaller than the value expected geometrically by approximately 10^{-8} ohm (i.e., by 25%). This decrease can be caused only with the healing of the micro-crack which gradually developes ahead of the pile-up of dislocations as the pile-up increases (the accumulation of an incomplete local shear), and which disappears together with the additional resistance contributed by it when the avalanche reaches the surface.[6] The quoted value of 10^{-8} ohm then corresponds to a crack measuring several microns.[46] An analysis of relation (12) shows that the crack must have contained under these conditions $n_s \sim$ 100 to 200 dislocations; the remaining (greater) part of the pile-up represents a "tail" which did not merge.

3. Since the resistance introduced by the crack $\sim c^2 \sim n_s^4$,[9] while n_s is determined in turn by the total number n of dislocations in the pile-up it is natural to expect the difference between the actual and the geometrically-expected increment of resistance to diminish rapidly with increasing avalanche. Indeed, in the case of elementary steps smaller than a unit step, experiment gives good support to this hypothesis.[6,58]

4. With very rare exceptions, oscillograms show no stepwise resistance increments not accompanied by deformation steps or considerably greater (apart from experimental error) than the values predicted geometrically. Consequently, apparently no microcracks of appreciable size developed suddenly in these experiments.[6]

5. The fraction of staircase deformation increases sharply when impurities are present in the crystal, sometimes reaching 100% (Fig. 21); consequently the presence of foreign atoms (point defects) in the crystal lattice appreciably aids the formation of dislocation avalanches.

If we now substitute into (20) the experimentally obtained values $n \sim 10^3$ and $\tau \sim 2 \times 10^7$ dynes/cm^2 (for average orientations $\chi \sim 45°$), we obtain for zinc $k_{Zn} \cong 0.6$ if $\sigma \sim 10^3$ erg/cm^2 and $b \sim 3 \times 10^{-8}$ cm.

If a pile-up of incomplete shears $\Delta = nb$ occurs in a plane oriented at an angle χ with respect to the applied tensile stress, the single crystal undergoes brittle fracture when n reaches, at a given τ_0 (or P), the "dangerous" value

$$N = \frac{k\sigma}{b\tau_0}\cot\chi = \frac{k\sigma}{\alpha P}\cdot\frac{1}{\sin^2\chi}; \quad k_{Zn} \sim 1. \tag{21}$$

However, whether under given conditions the crystal will in fact be brittle or plastic will depend upon the values that can be attained by the incomplete local shears, i.e., upon the actual maximum number n of dislocations in the pile-up.

One can expect that, for small shear stresses (up to the plastic flow limit), n will be determined essentially by the "effective" size of the slip plane L, in accordance with relation (7):

$$n^{(1)} \simeq \varkappa\frac{L\tau_0}{Gb}; \tag{22}$$

since experiment gives $n \sim 10^3$ for $L \sim 0.1$ cm and $\tau_0 \sim 100$ to 200 g/mm^2, the dimensionless factor κ turns out to be somewhat larger than given by (7).

For large shear stresses the size of the pile-ups is determined, apparently, by the nature of the

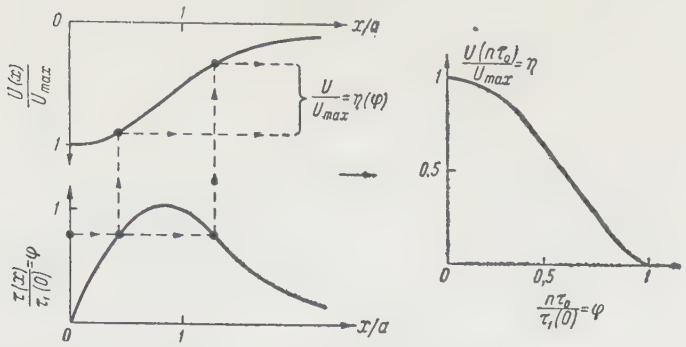


FIG. 22. Method of computing the function $\eta(\varphi)$ from the known function U(x).

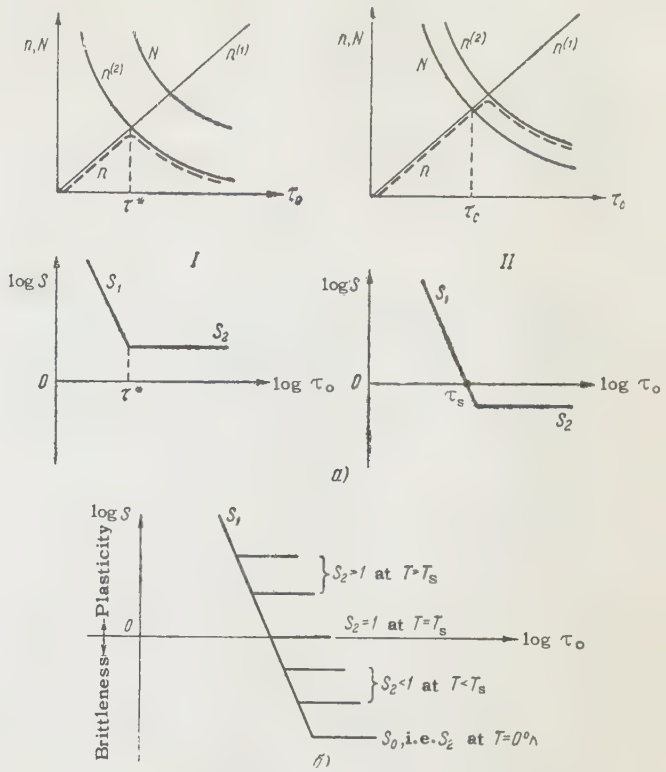


FIG. 23. (a) Dependence of the maximum number n of dislocations in a pile-up, of the "dangerous" number of dislocations N, and of the branches of the deformability criterion S_1 and S_2 on the applied shearing stress τ_0: I) $n^{(2)} < N$ and $S_2 > 1$; the crystal is plastic, τ^* should be interpreted as the plastic flow limit; II) $n^{(2)} > N$ and $S_2 < 1$; the crystal is brittle, τ_s is the fracture strength. (b) Temperature dependence of the branch S_2 of the deformability criterion.

obstacles encountered by the dislocations in the slip plane, specifically by the distribution of obstacles and the ease with which they can be overcome. In single crystals of pure metals the resistance to shear is determined in particular by the interaction of parallel dislocations lying in neighboring slip planes, by stationary dislocations, by the interaction of intersecting dislocations, and by the appearance of jogs when they cross, jogs which may give rise during subsequent motion to the appearance of chains of vacancies or of interstitials, etc. In polycrystalline samples the decisive role is played by the grain boundaries. In alloys the dislocations must overcome in the course of their motion inclusions of foreign atoms with some degree of dispersion; an obstacle to motion is also provided by adsorption on the dislocations of solute atoms ("Cottrell atmosphere").[61-63] Of particular interest to us is the interaction of dislocations with the free surface of the crystal.

The first very rough approximation[64] consists of an investigation of obstacles of only one type, those which are most significant under given conditions and which are characterized by a potential barrier $U(x)$. The height of the barrier U_{max} (erg/atom) determines the activation energy for overcoming the obstacle at $\tau_0 = 0$, while the quantity $|\text{grad } U(x)|_{max}/b^2 = \tau_1(0)$ is the local shearing stress required for breakthrough in the absence of thermal activation, i.e., at $T = 0°K$.[61,65] If the pile-up contains n dislocations, the force per unit length of the leading dislocation is $nb\tau_0$;[38] in this case the activation energy is $U(n\tau_0) = U_{max} \cdot \eta(\varphi)$, where $\varphi = n\tau_0/\tau_1(0)$; the function $\eta(\varphi)$ is determined by the shape of the barrier $U(x)$ (Fig. 22). (The parameter a in Fig. 22 characterizes the extension of the obstacle in the direction of motion of the dislocation and is close in order of magnitude to atomic distances). Assume that v dislocations arrive at the obstacle each second. Then the rate of their pile-up $\dot{n} = v - \nu\lambda \exp[-U(n\tau_0)/kT]$, where ν is the frequency of lattice oscillations and λ is the number of points at which a breakthrough is possible (it is assumed that to release the dislocations it is sufficient to have a breakthrough at one spot[45]). Dynamic equilibrium corresponds to $\dot{n} = 0$; consequently, for fairly large values of τ_0, when the obstacles become surmountable, the maximum number of dislocations in a pile-up is determined by the relation $n^{(2)} = \tau_1/\tau_0$, where the local stress is $\tau_1 = \tau_1(0)\varphi(\eta)$ and $\eta = kT \ln(\nu\lambda/v)/U_{max}$.

The parameter v introduced by us is directly related to the rate of shear $\dot{\varepsilon}$ and to the microinhomogeneity of the deformation (localization of shears along the slip lines), including the average thickness of slip bands h. The shear $\dot{\varepsilon}$ referred to h cm and measured in units of b, i.e., the quantity $v_0 \sim \dot{\varepsilon}h/b$ gives approximately the average number of dislocations crossing the crystal per second, within the limits of one slip line. The decisive role is played, however, not by this average value but by its maximum value $v = v_0 u$ where the deformation micro-inhomogeneity factor u has the meaning of the ratio of the total number of

slip lines to the minimum number of active lines. If we note that for average orientations of the crystal χ the rate of shear $\dot{a}$ is close in order of magnitude to the rate of elongation $\dot{\epsilon}$, we obtain for the parameter v the following approximate estimate:

$$v \sim \frac{\dot{\varepsilon} hu}{b}. \qquad (23)$$

When $\dot{\epsilon} \sim 10\%\ \text{min}^{-1}$ and the deformation has the greatest possible homogeneity on the microscale, i.e., when $h \sim 10^{-4}$ cm and $u = 1$, the rate of local slipping amounts to $v_0 \leq 10$ translational units per second. The second extreme value of v occurs when $hu \sim 1$ cm and (within the limits of macroscopically smooth flow at a constant rate, $\sim 10\%\ \text{min}^{-1}$) attains a value $v_{max} \sim 10^5\ \text{sec}^{-1}$; also of the same order of magnitude is the rate of slipping for the elementary deformation jumps[6,57] which we have observed and which produce a shear of 500 b to 1000 b within a time of the order of several milliseconds, at small average rates of elongation $\dot{\epsilon} \sim 0.01 - 1.0\%\ \text{min}^{-1}$.

For the true maximum value of the number of dislocations in a pile-up it is evidently possible to take $n = \min\{n^{(1)}, n^{(2)}\}$. The condition $n < N$ corresponds to dislocations overcoming the barrier and to plastic flow; for $n > N$ a nonequilibrium crack is developed, meaning that brittle fracture occurs. Therefore the value $S = N/n$ has the meaning of a criterion for the deformability of a crystal.[64] By making use of the values of N, $n^{(1)}$, and $n^{(2)}$ found above we can write the criterion S in the following form:

$$S \begin{cases} S_1 = A \frac{\sigma G}{L} \cdot \frac{1}{\tau_0^2} \cdot \cot\chi & \text{for } \tau_0 < \tau^*, \\ S_2 = B \frac{\sigma}{b\tau_1} \cdot \cot\chi & \text{for } \tau_0 > \tau^*, \end{cases} \qquad \tau^* = C\sqrt{\frac{Gb\tau_1}{L}}, \qquad (24)$$

$$\tau_1 = \tau_1(0)\,\varphi(\eta) = \tau_1(0)\,\varphi\left(\frac{kT}{U_{max}} \ln\frac{\nu\lambda b}{\dot{\varepsilon} hu}\right), \qquad (25)$$

where $A = k/\kappa$, $B = k$, and $C = 1/\sqrt{\kappa}$ are dimensionless coefficients close to unity in order of magnitude; for a qualitative discussion they can be omitted.[64]

If $S_2 < 1$, brittle fracture occurs at $S_1 = 1$, i.e., at a stress $\tau_S = \sqrt{A(\sigma G/L)\cot\chi}$, which rigorously corresponds to Eq. (17); we note that $k/\kappa^2 = \gamma^2$. If $S_2 > 1$, then τ^* has the meaning of the plastic flow limit. The characteristic dependence of the plastic flow limit on the grain size L (in the case of polycrystalline samples) has been well substantiated by experiment.[45]

According to (24) and (25) the critical temperature T_c of the transition from brittleness to plasticity is equal to[6,64]

$$T_s = \frac{U_{max}\eta(S_0)}{k \ln\frac{\nu\lambda b}{\dot{\varepsilon} hu}}, \text{ where } S_0 = \frac{\sigma}{b\tau_1(0)}. \qquad (26)$$

In accordance with the definition of v, the "probability" of one leading dislocation becoming free during the time 1/v, i.e., during the time between the arrivals of successive dislocations from the source, is given for $\tau_0 = \tau_S$ by

$$M = \frac{\nu\lambda}{v} e^{-\frac{U_{max}\eta(S_0)}{kT}} = e^{-\frac{U_{max}\eta(S_0)}{k}\left(\frac{1}{T}-\frac{1}{T_c}\right)}. \qquad (27)$$

At $T = T_S$ this quantity is equal to 1, for $T < T_S$ it rapidly tends to 0, while for $T > T_S$ it increases sharply. The quantity M can be called the "measure of plasticity" of the solid. However, for comparison with experiment it is more convenient to redefine this quantity in such a way that it gives a value of unity for $T \to \infty$; this result will be obtained if we regard the function

$$W = 1 - e^{-M}. \qquad (28)$$

as the "probability of plastic flow."

A completely analogous expression can be obtained, according to Stroh,[45] in the following manner. Let Λ be the length of a fixed segment of a dislocation line. Then, assuming that this segment is held fixed by a condensed atmosphere of interstitial atoms,[65] we obtain for the possible number of points at which a breakthrough may occur $\lambda = \Lambda/b$. Further, let U be the activation energy of the process at a stress τ_0. Then the probability of a breakthrough of a dislocation past the obstacle during a time dt is equal to $w = \nu\lambda e^{-U/kT}\,dt$.

If p is the probability that the breakthrough has not yet occurred at the time t, the probability of its occurring during the time interval from t to t + dt is equal to the product pw of the probabilities. According to the definition of p, it gives the decrement in p during the given time interval: $p\nu\lambda e^{-U/kT}\,dt = -dp$. Integration of this equation yields

$$p = \exp\left\{-\int_{-\infty}^{t} \nu\lambda e^{-U/kT}\,dt\right\}.$$

By assuming that U is approximately independent of t and on taking for our range of integration a certain value t_1 (Stroh[45] interprets it as the "time required for a crack to occur," $\sim 10^{-6}$ sec), we have for the probability of plastic flow: $W = 1 - p = 1 - e^{-\nu\lambda t_1 e^{-U/kT}}$, which coincides with (28) for $t_1 = 1/v$.

In accordance with (28), the quantity W varies from 0.1 to 0.9 within the limits of the following

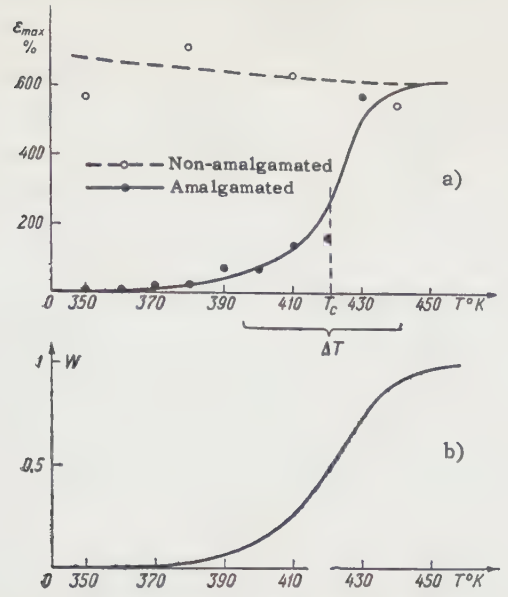


FIG. 24. (a) Dependence of the limiting deformation preceding fracture ε_{max} on the temperature for pure and amalgamated zinc single crystals at $\dot{\varepsilon} \cong 15\%\ min^{-1}$ and $\chi \cong 45°$. (b) Theoretically calculated temperature dependence of the "probability of plastic flow" W for amalgamated single crystals.[6,21]

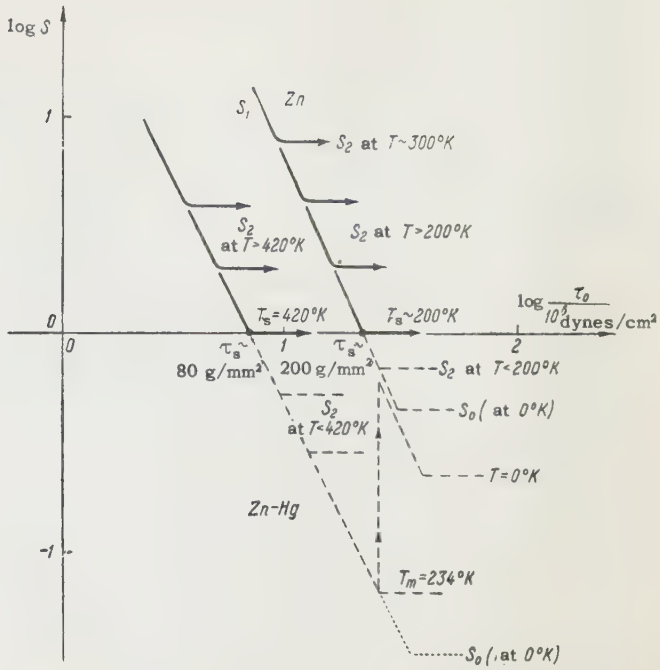


FIG. 25. The deformability criterion S for pure and amalgamated zinc single crystals at $\dot{\varepsilon} \sim 10\%\ min^{-1}$, $L \sim 0.1$ cm and $\chi \sim 45°$.

temperature interval situated in the neighborhood of T_S

$$\Delta T \approx 3.3 \frac{kT_s^2}{U_{max}\,\eta(S_0)}. \qquad (29)$$

In our case of brittle fracture of amalgamated single crystals of zinc (Fig. 24), $T_S \cong 420°K$, while the transition interval ΔT amount to 40 — 50°; correspondingly $U_{max}\,\eta(S_0) = 1.1 \pm 0.1$ ev. (Since σ is small in this case, S_0 is considerably less than 1, while $\eta(S_0) \cong 1$, i.e., U_{max} has approximately the same value, 1.1 or 1.2 ev.) The relatively small critical value which appears in the logarithm of (26) can be estimated as follows:

$$\ln \frac{\nu\lambda b}{\dot{\varepsilon} h u} \approx 3.3 \frac{T_s}{\Delta T}, \qquad (30)$$

in this case $3.3\,T_S/\Delta T = 31 \pm 3$. The function $W(T)$ which corresponds to these estimates is plotted below in Fig. 24 and agrees completely with the experimentally determined function form of $\epsilon_{max}(T)$.

Figure 25 gives a qualitative idea of the distribution of the branches of the criterion S for pure and amalgamated single crystals of zinc corresponding to $\dot{\varepsilon} \sim 10\%\ min^{-1}$, $L \sim 0.1$ cm, and $\chi \sim 45°$. Crystals of pure zinc exhibit brittleness at low temperatures; at ordinary temperatures they are highly plastic, and their plastic flow limit is not very high. In the case of crystals coated with mercury σ decreases sharply not only on the surface of the crystal, but also on the internal micro-surfaces, owing to the rapid two-dimensional migration of mercury atoms along hollow dislocation nuclei and along the resultant micro-cracks. The straight lines S_1 and S_2 are displaced downwards; a further lowering of S_2 may be due to a higher value of U_{max} for amalgamated crystals. The new position of the branches of S indicates catastrophic brittleness with a considerable decrease of strength up to $T_S = 420°K$; a high degree of plasticity is re-established for $T > T_S$. Below $T_m = 234°K$, mercury stops migrating and consequently σ stops and ordinary strength is re-established. As a result of a discontinuous jump S_1 becomes coincident with the line S_1 for pure crystals; only a difference in the temperature dependence of S_2 may remain, corresponding to a high degree of hardening of crystals with interstitial foreign atoms.[23]

If the melting point of the original metal is T_M, while that of some low melting point component (or coating) capable of producing a considerable decrease of σ is T_m, the following three cases are possible:[64]

(a) $T_m < T_S < T_M$ — brittleness accompanied by decreased strength in a restricted temperature range $T_m - T_C$ (for example, Zn-Hg[21]);

(b) $T_m < T_M \lesssim T_S$ — brittleness accompanied by diminished strength from T_m up to T_M (for example, Zn-Sn[18]);

(c) $T_S < T_m$ — absence of an effect (condition for the suitability of a given additive).

A completely different picture presents itself for weak (organic) adsorption-active media, which do not make the crystal brittle, but which are capable under certain conditions of making it plastic.[3-5,64] A small change in σ cannot by itself displace noticeably the branches of S. However, if under given conditions of deformation the interaction between the dislocations and the free surface of the crystal plays an essential role, even a small decrease in σ can raise the branch S_2 appreciably and thereby displace to the left the plastic flow limit point τ^*.

The corresponding energy barrier, $U_{max} = b^2\sigma = 0.5$ ev, is due to the formation of a step at the place where the dislocation reaches the surface and is quite steep, for example, $U(x) \sim -U_{max} \times [1 + (x/a)^3]^{-1}$ (a is of the order of atomic distances), i.e., in the immediate neighborhood of the surface the force determined by it predominates over the slowly varying repulsive "image force."[61] (It is not excluded that a still more important role is played by surface defects which block individual points on the trace of the slip plane on the surface; in this case the barrier may be higher and may have another form, but as before it retains its short-range character and depends essentially on σ.)

Then, in accordance with (24) and (25), we have the following expression for the decrease in the plastic flow limit as a function of σ:[64]

$$\frac{\Delta\tau_\sigma^*}{\tau^*} = \frac{1}{2}\left(1 + \frac{\tau_i}{\varphi}\left|\frac{\partial\varphi}{\partial\tau_i}\right|\right)\frac{\Delta\sigma}{\sigma}. \qquad (31)$$

Equation (31) means that when $\eta \sim 0.9$ a decrease in σ of only a few tens of ergs/cm^2 should reduce the plastic flow limit by a factor of two. The optimum in the plasticizing action of the adsorption-active medium at a given rate of deformation corresponds, from this point of view, to the value $\eta \sim 1$, i.e., the upper temperature limit of the effect of the above potential barrier is given by:

$$T_{opt} \sim \frac{U_{max}}{k \ln \frac{\nu\lambda b}{\dot{\varepsilon} h a}}. \qquad (32)$$

The dependence of the optimum in the effect of the adsorption-active medium on the temperature and on the rate of deformation were studied by Likhtman et al.[66] They found $T_{opt} \sim 290°K$ for $\dot{\epsilon} \sim 10\%\ min^{-1}$, and $T_{opt} \sim 370°K$ for $\dot{\epsilon} \sim 500\%\ min^{-1}$. In accordance with (32), we obtain from this $U_{max} \sim 0.5$ ev, which is in full agreement with our assumption concerning the nature of the obstacle that gives rise to the plasticizing effect. The logarithmic term equals approximately 20 (in this model the dislocation is fixed only at the end that emerges on the surface, i.e., $\lambda \sim 1$ is much smaller than in the case discussed above).

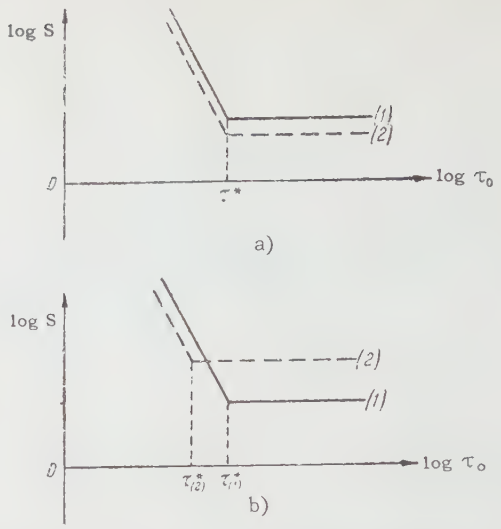


FIG. 26. (a) Change in S due to a small change in σ, if U does not depend on σ. (b) A possible change in S corresponding to a small change in σ, if $U \propto \sigma$.

Nevertheless, when $\eta \sim 1$ expression (31) cannot, generally speaking, give a quantitative estimate of the effect, since in this case only the surface energy barrier is taken into account. The simplest method of taking into account the superposition of hardening factors in this model is to replace τ_1 [cf. (24)] by the sum $\tau_1 + \tau_i$, where τ_i takes into account the remaining hardening factors and depends in the first approximation only on the magnitude of the deformation ϵ (shear a), but not on the temperature.

Applications of the model described above are closely associated with a careful analysis of the temperature dependence of the hardening factors and of the spectra of their activation energies. Calculation of local stresses $\tau_1(T)$ due to overcoming barriers requires, like in references 61 and 65, the introduction of specific models of interaction between dislocations and obstacles. But even a qualitative investigation along these lines leads to a number of interesting conclusions.

Thus, it is natural to expect that when $\dot{\epsilon}$ is reduced by several orders of magnitude the critical temperature T_S for the transition from brittleness to plasticity drops, in accordance with (26), sufficiently to eliminate brittleness produced by a low-melting-point coating. Moreover, in accordance with (31), there even exists in this case the possibility of plasticizing, a possibility previously masked by the embrittlement due to a sharp decrease of σ on the incipient micro-cavities and,

to a certain extent, by the effect of alloying. Indeed, the experiments of reference 17 show convincingly that at creep rates $\dot{\epsilon} \sim 10^{-4}$ % min^{-1}, coating of single crystals of zinc with molten tin increases the rate of steady creep by a factor of approximately two, while at $\dot{\epsilon} \sim 10\%\ min^{-1}$ the same coating results in brittleness accompanied by sharply reduced strength.

However, it is necessary to point out a basic limitation on the applicability of the above approximate method of deriving the criterion of deformability. This limitation is determined, apparently, not so much by taking into account only one type of potential barrier as by identifying the obstacle in front of which the "dangerous" accumulation is formed, and the obstacles that determine the plastic flow limit. In particular, the plastic flow limit turns out in this case to depend essentially on the process of formation of pile-ups of dislocations. Such an approach is valid if the predominant role in both processes is actually played by obstacles of one type (for example, grain boundaries in polycrystalline samples); in other cases it may turn out to be, generally speaking, insufficient.

Indeed, let us examine the following possible case. Suppose that at least one active slip plane contains a very strong obstacle (practically insurmountable), in front of which an accumulation of dislocations is formed, while obstacles to shear scattered in other glide planes may be overcome even at not very large values of τ_0, for example, by single dislocations. In such a case, the branch S_2 loses its significance, while the first branch assumes the form:

$$S_1 = A \frac{\sigma G}{L} \frac{1}{\tau^2(\dot{\epsilon}, T, a, \ldots)} \cot \chi, \qquad (33)$$

where $\tau(\dot{\epsilon}, T, a, \ldots)$ is the plastic flow stress determined by the total crystal volume and, in particular, by the mechanical hardening due to an increase in the shear a. Interpreted in this way the quantity

$$S_1^{\frac{1}{2}} = \frac{K\sqrt{\cot \chi}}{\tau} = \frac{\tau_0}{\tau} \qquad (34)$$

is, evidently, the ratio of the "dangerous" shearing stress determined by the condition (17) to the value of the actual shearing stress at a given stage of deformation and under given conditions (temperature, rate of elongation). In this latter form the criterion S corresponds to an investigation of the transition from brittleness to plasticity in accordance with the well known method proposed by A. F. Ioffe.[67]

If the temperature dependences of the plastic flow limit and of the hardening coefficient are known expression (34) allows us, obviously, to calculate the limiting deformation ϵ_{max} and the fracture stress P_S for a single crystal with arbitrary initial orientation χ_0, since the glide plane and the cleavage plane coincide.

But if the dangerous crack develops at an angle Θ to the glide plane, then the factor $\sqrt{\cot \chi}$ should be replaced by the appropriate more complicated function $f(\chi, \Theta)$.

[1] P. A. Rebinder, Z. Physik **72**, 191 (1931); Юбилейный сборник, посвящ. 30-летию Великой Октябрьской Социалистической революции. Изд. (Commemorative Collection on the 30-th Anniversary of the Great October Socialist Revolution) Acad. Sci. Press, **1**, 123 (1947).

[2] P. A. Rebinder, Izv. Akad. Nauk S.S.S.R., OKhN (Div. Chem. Sci.) **11**, 1284 (1957).

[3] V. I. Likhtman and P. A. Rebinder, Izv. Akad. Nauk S.S.S.R., Ser. Fiz. **17**, 313 (1953).

[4] V. I. Likhtman, Usp. Fiz. Nauk, **37**, 3 (1949); **54**, 6 (1954).

[5] Likhtman, Rebinder, and Karpenko, Влияние поверхностно-активной среды на процессы деформации металлов (Effect of a Surface-Active Medium on the Processes of Deformation of Metals) Acad. Sci. Press, 1954.

[6] E. D. Shchukin, A Study of the Inhomogeneities in the Plastic Deformation of Metallic Single Crystals, Thesis, Inst. of Phys. Chem., Acad. Sci. U.S.S.R., Moscow, 1957.

[7] C. Benedicks, Compt. rend. **232**, 2435; **233**, 409, 482 (1951); M. Sato, Proc. Jap. Acad. **30**, 193, 369 (1954).

[8] M. Volmer, Z. phys. Chem. **125**, 151 (1927); **A155**, 281 (1931).

[9] P. A. Rebinder, New Problems of Physico-Chemical Mechanics, Report at the Permanent Colloquium on Solid Phases of Variable Composition, held jointly with the Moscow Colloid Colloquium, Jan. 26, 1956.

[10] W. Goodrich, J. Iron. Steel Inst. **132**, 43 (1935).

[11] R. Genders, J. Inst. Met. **37**, 215 (1927).

[12] G. Austin, J. Inst. Met. **58**, 173 (1936).

[13] L. Ewijk, J. Inst. Met. **56**, 241 (1935).

[14] H. Miller, J. Inst. Met. **37**, 183 (1927).

[15] E. Leech, P. Gregory, R. Eborall, J. Inst. Met. **83**, 347 (1954-55).

[16] Kishkin, Nikolenko, and Ratner, J. Tech. Phys. (U.S.S.R.) **24**, 8, 1455 (1954); Ya. M. Potak and I. M. Shcheglakov, J. Tech. Phys. (U.S.S.R.) **25**, 5, 897 (1955).

[17] V. A. Labzin and V. I. Likhtman, Dokl. Akad. Nauk S.S.S.R. **121**, 778 (1958), Soviet Phys. "Doklady" **3**, 443 (1958).

[18] Rebinder, Likhtman, and Kochanova, Dokl. Akad. Nauk, S.S.S.R. **111**, 1278 (1956).

[19] Likhtman, Kochanova, and Bryukhanova, Dokl. Akad. Nauk S.S.S.R. **120**, 757 (1958), Soviet Phys. "Doklady" **3**, 599 (1958).

[20] V. I. Likhtman and B. M. Maslennikov, Dokl. Akad. Nauk S.S.S.R. **67**, 1 (1949).

[21] Rozhanskii, Pertsov, Shchukin, and Rebinder, Dokl. Akad. Nauk S.S.S.R. **116**, 769 (1957), Soviet Phys. "Doklady" **2**, 460 (1957).

[22] E. Schmid and V. Boas, Plasticity of Crystals, (Russian translation), GONTI, 1938

[23] L. A. Bryukhanova and E. D. Shchukin, Инженерно-физический журнал (Eng.-Phys. J.) **1**, No. 8, 116 (1958).

[24] E. D. Shchukin and V. I. Likhtman, Dokl. Akad. Nauk S.S.S.R. (1958) (in press).

[25] M. Polanyi, Z. Physik **7**, 323 (1921).

[26] F. Zwicki, Z. Physik **24**, 131 (1923).

[27] C. Zener, Elasticity and Anelasticity of Metals, Chicago (1948).

[28] A. A. Griffith, Phil. Trans. Roy. Soc. **A221**, 163 (1920).

[29] R. A. Sack, Proc. Phys. Soc. **58**, 729 (1946).

[30] E. Orowan, Z. Krist. **A89**, 327 (1934).

[31] H. A. Elliot, Proc. Phys. Soc. **59**, 208 (1947).

[32] G. Neiber, Концентрация напряжений (Stress Concentration), Gostekhizdat (1947).

[33] N. F. Mott, J. Phys. Soc., Japan **10**, 650 (1955).

[34] N. F. Mott, Proc. Phys. Soc. **B64**, 729 (1951).

[35] N. F. Mott, Phil. Mag. **43**, 1151 (1952).

[36] A. H. A. Mèlèka, Phil. Mag. **1**, 803 (1956).

[37] J. D. Eshelby, F. C. Frank, F. R. N. Naharro, Phil. Mag. **42**, 351 (1951).

[38] J. S. Koehler, Phys. Rev. **60**, 397 (1941).

[39] V. L. Indenbom and G. E. Tomilovskii, Кристаллография (Crystallography) **2**, 190 (1957).

[40] V. L. Indenbom and M. A. Chernysheva, Dokl. Akad. Nauk S.S.S.R. **111**, 596 (1956).

[41] J. S. Koehler, Phys. Rev. **85**, 480 (1952).

[42] A. N. Stroh, Proc. Roy. Soc. **A223**, 404 (1954).

[43] N. F. Mott, Proc. Roy. Soc. **A220**, 1 (1953).

[44] A. N. Stroh, Proc. Roy. Soc. **A218**, 391 (1953).

[45] A. N. Stroh, Phil. Mag. **46**, 968 (1955).

[46] A. N. Stroh, Phil. Mag. **2**, 1 (1957).

[47] F. E. Fujita, J. Phys. Soc., Japan. **11**, 1201 (1956).

[48] J. Fisher, Acta Met. **3**, 109 (1955).

[49] R. Green, Phys. Rev. **102**, 376 (1956).

[50] F. Seitz, Adv. in Phys. **1**, 43 (1952).

[51] L. M. Clarebrough, M. E. Hargreavs, G. W. West, Proc. Roy. Soc. **A232**, 252 (1955); Phil. Mag. **1**, 528 (1956).

[52] N. J. Petch, Phil. Mag. **1**, 331 (1956).

[53] E. Orowan, Nature **154**, 341 (1944).

[54] N. J. Petch, Progr. in Met. Phys. **5**, 1 (1954).

[55] R. Peierls, Proc. Phys. Soc., **A52**, 134 (1940).

[56] Ya. I. Frenkel', Z. Physik **37**, 572 (1926).

[57] Shchukin, Goryunov, Pertsov and Rozhanskii, Dokl. Akad. Nauk S.S.S.R. **118**, 277 (1958), Soviet Phys. "Doklady" **3**, 96 (1958).

[58] Shchukin, Rozhanskii, and Goryunov, Dokl. Akad. Nauk S.S.S.R. **115**, 1101 (1957), Soviet Phys. "Doklady" **2**, 420 (1957).

[59] Shchukin, Pertsov and Rozhanskii, Приборы и техника эксперимента (Instr. and Exp. Eng.) **2**, 98 (1957).

[60] D. Kuhlmann, Z. Metalkunde **41**, 129 (1950).

[61] A. H. Cottrell, Dislocations and Plastic Flow in Crystals (Russian translation), Metallurgizdat, 1958.

[62] G. Schoeck, Adv. in Appl. Mech. **4**, 229 (1956).

[63] A. Seeger, Z. Naturforsch. **9a**, 758, 856, 870 (1954).

[64] E. D. Shchukin, Dokl. Akad. Nauk S.S.S.R. **118**, 1105 (1958), Soviet Phys. "Doklady" **3**, 143 (1958).

[65] A. H. Cottrell, B. A. Bilby, Proc. Phys. Soc. **A62**, 49 (1949).

[66] Likhtman, Rebinder, and Yanova, Dokl. Akad. Nauk S.S.S.R. **56**, 8 (1947).

[67] A. F. Ioffe, Физика кристаллов (Crystal Physics) Gosizdat, Moscow, 1929.

[68] E. D. Schchukin and P. A. Rebinder, IV Всесоюзная конференция по коллоидной химии (IV All-Union Conference on Colloid Chemistry) (Tbilisi, May 1958), Abstracts of Papers, Acad. Sci. Press, Moscow, 1958, p. 128.

[69] V. N. Rozhanskii, Usp. Fiz. Nauk **65**, 387 (1958).

Translated by G. Volkoff

DETECTION OF ION BEAMS BY ELECTRON MULTIPLIERS IN MASS SPECTROMETRY AND NUCLEAR PHYSICS

A. I. AKISHIN

Usp. Fiz. Nauk **66**, 331-346 (October, 1958)

INTRODUCTION

In recent years there has been an increasing use of mass spectrometry in the analysis of radioactive fuel products and in the search for traces of nuclides that are formed in nuclear reactions; in this connection the need has arisen for a very sensitive ion detector. The development of magnetless radio-frequency mass spectrometers requires simple ion detectors with no time delay. The use, in these applications, of the usual detection schemes, which use electrometers or vacuum-tube amplifiers, involves a number of difficulties when one wishes to detect ion currents smaller than 10^{-15} to 10^{-16} amp. It is more convenient to use electron multipliers as detectors; these multipliers have alloy cathodes with which it is possible to detect charged particle beams characterized by intensities of 10^{-19} amp and pulse lengths of approximately 10^{-9} sec. In nuclear physics these multipliers are frequently used for the detection of recoil nuclei with energies of several tens or hundreds of electron volts, that is to say, when the use of other types of counters is not feasible.

In the present article we review the transformation of the ion beam into an electron beam at the cathode of the multiplier, the efficiency of the multiplier for counting particles, and the gain stability of a multiplier. Various examples are considered, and certain features of the application of electron multipliers as ion detectors are considered in connection with mass spectrometry and nuclear physics. A great deal of important information concerning the characteristics and construction of electron multipliers can be found, for example, in Electron Multipliers by N. O. Chenka, S. M. Fainshtein and T. M. Lifshitz.

1. TRANSPORTATION OF AN ION BEAM INTO AN ELECTRON BEAM

A photograph of an electron multiplier with the cover removed is shown in Fig. 1. Multipliers of this type are used frequently for the detection of ion beams. The transformation of the ion beam into an electron beam takes place at the cathode of the multiplier as the result of the ejection of secondary electron ions. By definition the emission coefficient (γ) is the number of electrons emitted by the target divided by the number of heavy particles which strike the target in the same time period. The value of γ is a linear function of ion energy up to several kilo electron volts; the linear region of $\gamma(E)$ increases as the mass number of the bombarding ions increases.[1,2]

Figure 2 shows the output current of an electron multiplier used to detect different ions; the current is plotted as a function of ion energy, which is increased from 5 to 250 kev.[3] The masses of the detected ions vary from 1 to 40. The multiplier has 11 activated beryllium-copper dynodes. The gain of the multiplier is approximately 10^6. From the behavior of the curve shown in Fig. 2 it is apparent that the output current from the multiplier is higher over the entire energy range for high-mass ions when the ion energy is greater than 15 to 25 kev. On the other hand, at ion energies below 15 to 25 kev the light-ion current is higher. For a proton beam the maximum output current (the maximum value of γ) is observed at approximately 100 kev.[5] When the energy is increased beyond several hundred kilo electron volts, γ falls off slowly for protons and deuterons.[4-7]

As the mass of the bombarding ion increases the position of the maximum on the $\gamma(E)$ curve is shifted toward the high-energy region.[8,10] Thus, when He^+ ions bombard a molybdenum target whose surface is not atomically pure the maximum of the function $\gamma(E)$ is observed in the region from 200 to 400 kev.[8] On the average, one α particle ejects approximately 10 electrons from a metal;[9,10,27] this relation holds over a wide energy range.

Because of the increasing use of electron multipliers in mass spectrometry it is of considerable interest to know the dependence of γ on the mass of the ion, m_i; however, this relation had not been established until recently. Figure 3 shows a family of $\gamma(m_i)$ curves obtained by bombarding an activated target of silver manganese with ions of the

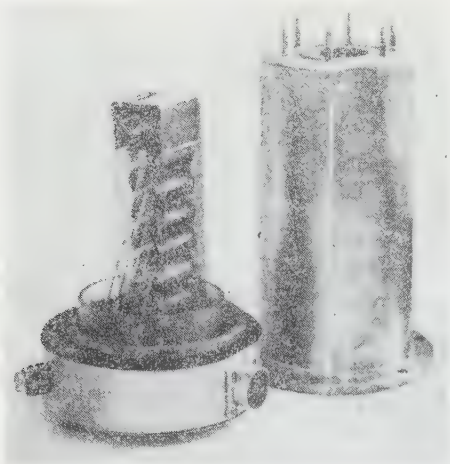


FIG. 1. Electron multiplier with the metal cover removed.

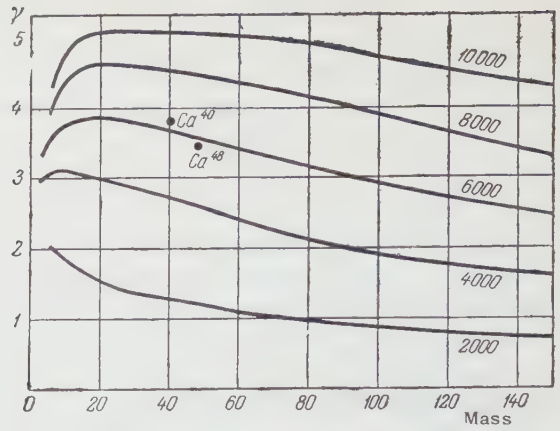


FIG. 3. Dependence of γ on ion mass for various values of the ion energy.

alkali metals at several values of ion energy.[2] At ion energies of approximately 2 kev, γ falls off with increasing ion mass. However, at ion energies of 8 to 10 kev there are sections on the $\gamma(m_i)$ curves where γ is more or less independent of ion mass. The same figure shows values of γ measured by bombarding targets with Ca^{40} and Ca^{48} ions with energies of 6 kev; it is apparent that $\gamma(m_i)$ varies more rapidly for ions of isotopes of the same element than for different elements. Apparently γ depends on the structure of the electron shells of the bombarding particles[1] as well as on the mass.

Ploch and Walcher[11-13] have investigated the function $\gamma(m_i)$ for the bombardment of Pt, Mo, Be, and Cu targets by the following positive ions: Li^6, Li^7, Ne^{20}, Ne^{22}, K^{40} and K^{41}; the surfaces of these targets were not atomically pure. These authors find that, as the ion energy is varied from 0.6 to 6 kev, ions of different isotopes (with the same velocity) eject the same number of secondary electrons.

It is found that molecular ions eject more electrons from metal targets than do atomic ions of the same energy, mass, and charge.[3,6,16] Figure 4 shows the output current of an electron multiplier with activated AgMg emitters (1.7% Mg) as a function of the mass and molecular structure of the detected ions.[1] The ion energy is 7 kev. The output current produced in detecting CH_2^+ ions is approximately 10% greater than that produced by atomic N^+ ions.

When targets are bombarded by molecular and atomic ions of higher mass, a considerable difference in the number of ejected electrons is observed. Thus, for example, 9-kev $C_{14}H_{10}^+$ ions eject approximately 2.2 times as many secondary electrons as do Hf_{178}^+ ions of the same energy;[15] these data refer to an electron multiplier with an activated beryllium-copper cathode.

The number of secondary electrons ejected from the metal depends on the magnitude and sign of the charge of the interacting particle. Negative ions give rise to more electrons than similar ions of positive charge;[1,17] neutral atoms produce approximately the same number of secondary electrons as singly-charged positive ions of the same energy.[18,19] A number of authors have observed that

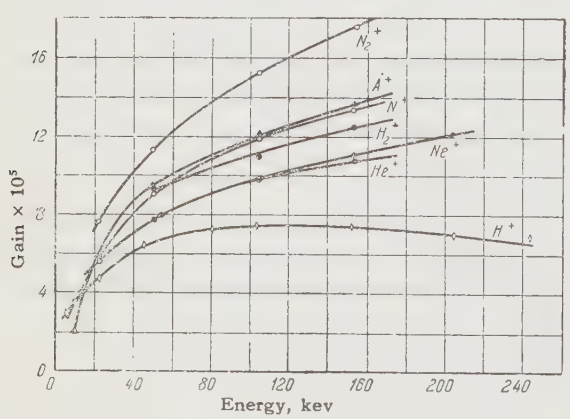


FIG. 2. Variation in output current of the multiplier as a function of ion energy for different ions.

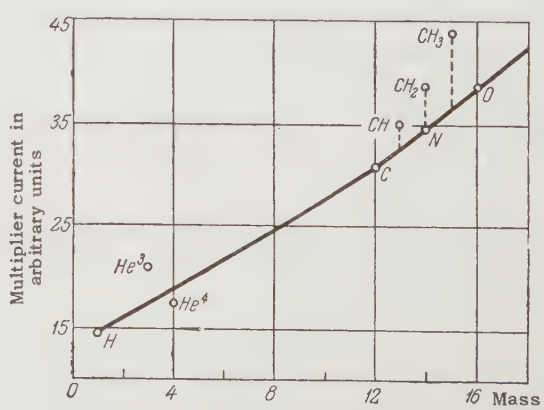


FIG. 4. Effect of the molecular configuration of the ion on the output current of the multiplier.

the value of γ increases as the charge of the positive ion increases.[20,22,23]

As the angle of incidence of the ion beam on the target increases the value of γ also increases.[1] For protons with energies of several hundred kev γ varies in proportion to the secant of the angle of incidence.[4,24]

The energy distribution of the secondary electrons is virtually unaffected by an increase in mass and energy of the bombarding ions. Most of the secondary electrons have energies below 30 ev.[8,21,22,24] The angular distribution of the secondary electrons is given roughly by the cosine law.[25]

2. COUNTING EFFICIENCY OF ELECTRON MULTIPLIERS FOR CHARGED PARTICLES

To determine the absolute intensity of an ion beam with an electron multiplier one must know the counting efficiency of the multiplier. In the ideal case each particle incident on the cathodes of the multiplier should give rise to a voltage pulse at the output; the amplitude of this pulse should be sufficient to ensure its being recorded by the counting circuit. However, since the electron emission due to the action of the bombarding particles is of statistical nature, there are always cases in which the incoming particle does not eject a secondary electron from the cathode. For this reason the counting efficiency of the electron multiplier is not exactly 100%.

The counting efficiency of a multiplier is usually determined by comparing the intensity of the particle beam incident on the cathode with the pulse counting rate in the counting circuit, as measured by means of an electrometer or a dc amplifier. In measurements of this kind the intensity of the particle beam is usually of the order of 10^{-15} amp (approximately 10^5 singly-charged particles per second); this guarantees a relatively small error when the ion beam is detected with an electrometer, dc amplifier, or counting circuit.

It has been shown experimentally[3] that H_1^+, H_2^+, He^+, Ne^+, M^+, A^+, and N_2^+ ions with energies from 10 to 200 kev are detected with an efficiency of $100^{+0\%}_{-5\%}$; these data refer to an 11-stage electron multiplier with an activated beryllium-copper emitter. In Fig. 5 the counting efficiency for N^+ ions is shown as a function of ion energy. The counting efficiency is 92% for an ion energy of 4 kev and almost 100% for ion energies between 10 and 200 kev. The reduction in counting efficiency for N^+ ions at 4 kev is apparently due to statistical fluctuations in γ. The mean value γ for bombardment of alloy emitters by N^+ ions with energies of 4 kev is approximately three.[1,2] If one assumes that the electron emission due to ions with energies of 4 kev has a Poisson distribution, the probability for events in which $\gamma = 0$ can be found from the expression $\beta_\gamma = (\bar{\gamma}^\gamma/\gamma!)\, e^{-\bar{\gamma}}$ where β_γ is the probability of ejection of γ secondary electrons by the ions. If $\gamma = 3$, then $\beta_{\gamma=0} \approx 0.05$ i.e., approximately 5% of the N^+ ions incident on the cathode do not eject secondary electrons.

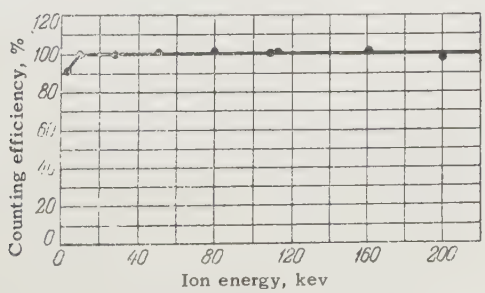


FIG. 5. Counting efficiency for N^+ ions as a function of energy.

In addition to the statistical fluctuations in γ, there is another effect that reduces the counting efficiency for 4-kev N^+ ions; this effect arises as the consequence of the wide amplitude distribution of the pulses from the multiplier. Some of the pulses are not recorded by the counting circuit because they are comparable in amplitude with the noise of the linear amplifier and lie below the amplitude threshold of the discriminator.

Morrish and Allen[26] have shown that electron multipliers with beryllium-copper emitters detect Li^+ ions with an efficiency of approximately 100% if the ion energy is greater than 2 kev. Allen[27] has noted that α particles with a wide energy distribution are detected by electron multipliers with an efficiency of about 100%. Neutral particles are detected with approximately the same efficiency as positive ions of the same energy.[3,28]

To obtain a high counting efficiency it is necessary that the sensitivity of the counting circuit be adjusted for the wide amplitude range of the pulses from the multiplier. The relative mean-square spread in M, the gain of an electron multiplier for ion detection, is given by the expression[27]

$$\frac{\overline{\Delta M^2}}{M^2} = \frac{\overline{\Delta\gamma^2}}{\gamma^2} + \frac{\Delta\sigma^2}{\gamma\sigma(\sigma-1)}, \qquad (1)$$

where σ is the mean gain per stage of the multiplier.

Equation (1) can be simplified by assuming that the emission of electrons due to ion collisions follows a Poisson distribution, i.e., $\overline{\Delta\gamma^2} = \gamma$, while $\overline{\Delta\sigma^2}$ is approximately 2σ for primary electron energies of approximately 400 ev for a beryllium emitter.[29] Equation (1) then becomes:

$$\frac{\overline{\Delta M^2}}{M^2} \approx \frac{1}{\gamma}\left[1+\frac{2}{(\sigma-1)}\right]. \qquad (2)$$

It follows from Eq. (2) that the relative mean-square spread M is inversely proportional to γ.

Electron multipliers with cathodes of large area are used for detecting wide ion beams or divergent ion beams. The counting efficiency for these multipliers also depends on the point of incidence on the cathode. Trebukhovskii et al.[31] describe a 15-stage electron multiplier with input-window dimensions 44×44 mm^2. The cathode and emitters of the multiplier are made of beryllium-copper. In this work the local characteristics of the multiplier were investigated by bombarding a cathode at a potential of approximately -5 kev with α particles from Po^{210}. It was found that the efficiency for counting α particles was reduced by edge effects at the cathode.

To obtain more uniform local characteristics the cathode of the multiplier must be kept at ground potential while the anode must be kept at a high positive potential. This mode of operation of the electron multiplier does not cause any significant reduction in the counting efficiency for ions if the initial ion energy is greater than several kev. When the initial energy of the ions is insufficient to cause detection by the electron multiplier, a high negative potential is applied to the cathode of the multiplier. In this case more uniform local characteristics can be obtained by using metal shields (which are kept at cathode potential) in the primary-electron region.

The counting efficiency for electrons and ions depends on the energy. The maximum electron counting efficiency is observed at electron energies of several hundred electron volts (Fig. 6);

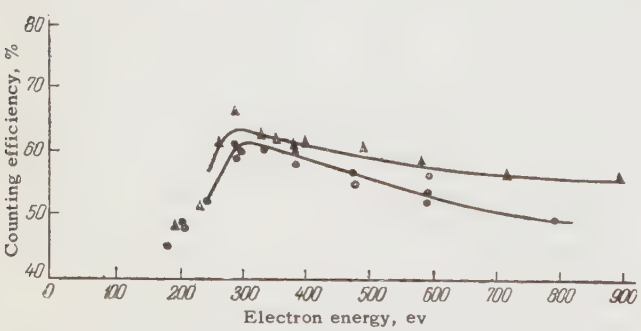


FIG. 6. Counting efficiency for N^+ ions as a function of energy. ▲) directly after activation of the multiplier, ●) after exposure of the multiplier to air for a period of five months.

this efficiency varies between 60 and 70%, depending on the degree of activation of the multiplier cathode.[32] According to the data reported by Allen[33] the counting efficiency for 500-ev electrons is almost 100%; at 6,000 ev this figure becomes 40%.

Electrons with energies of 30 kev are detected with an efficiency of approximately 23% by an RCA electron multiplier with beryllium-copper emitters. Beta particles with energies of 1 or 2 Mev are detected with efficiencies of several percent in electron multipliers.[34] The detection efficiency for electromagnetic radiation is several tenths of a percent.[35-36]

3. STABILITY OF MULTIPLIER GAIN

When electron multipliers are used in mass spectrometers or other physical research apparatus the electrodes are affected by various gases, ion bombardment, and so on. It is of interest to examine these effects, chiefly in connection with the gain stability of the electron multiplier.

Ion beams are most frequently detected by electron multipliers with beryllium-copper emitters which are relatively insensitive to exposure to air. After activation, a 12-stage electron multiplier of this kind was found to maintain a stable gain of the order of 10^6 for a period of several months even though air was allowed into the chamber.[30] Emitters of electron multipliers are also made from manganese alloys of copper, silver and aluminum;[1,36] however, these multipliers are more sensitive to air than those with emitters made from beryllium alloys.[38] To increase gain stability multipliers are sometimes made of unactivated beryllium-copper or of brass.[40] The gain of an 11-stage electron multiplier with slightly activated beryllium-copper emitters was approximately 10^5 for an interstage voltage of 400 volts.[3]

Fatigue effects are sometimes observed in the last emitters when intense ion beams are detected by electron multipliers; these effects cause a reduction in gain. Beryllium emitters are found more stable than magnesium in this respect.[38] Fatigue effects are observed in alloy emitters when the output current becomes somewhat greater than several milliamperes per square centimeter.

When several ion beams whose intensities differ by several orders of magnitude are detected by an electron multiplier within a short time period (for example, in using a mass-spectrometer to determine relative isotopic compositions) it is observed that the gain of the multiplier does not remain constant.[41] This instability is due primarily to charging of the insulators located near the trajectories of scattered electrons. To reduce this effect it is desirable to use an electrode system in which there is very little electron scattering and in which the number of insulators located near the trajectories is kept to a minimum. An electrode system of this type has been described in reference 42.

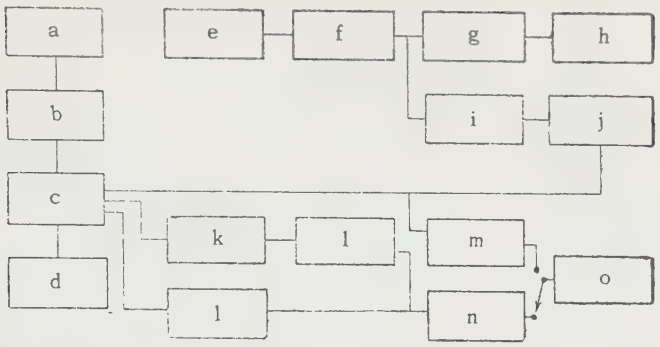


FIG. 7. Block diagram of the detection system in a mass spectrometer. a) Ion source; b) Fil. supply; c) Switch; d) Accel. volt. supply; e) 6,000-v power supply; f) Electron multiplier; g) Dynamic electrometer; h) Recorder; i) Preamplifier; j) Amplifier and discriminator; k) Scaler (10^2); l) Scaler (10^4); m) Integrator; n) Output meter; o) Recorder.

If the interstage voltages are low the gain stability is sensitive to variations in the voltage from the high-voltage power supply to the voltage divider. However, at interstage voltages of 500 to 600 volts the stability of the multiplier is much better.[3]

Transient discharges in multipliers reduce the gain by several percent.[3] If the discharge lasts for a minute or more the gain can be reduced by one or two orders of magnitude until the electrode system recovers. No change in multiplier gain is observed when the vacuum in the chamber deteriorates from 2×10^{-5} to 6×10^{-5} mm Hg.[3]

To obtain high stability the multiplier is generally mounted inside a metal chamber. In physical research electron multipliers are generally used as dc amplifiers or as particle counters. From the point of view of stability the latter type of operation is to be preferred. When a multiplier is used to count individual particles the threshold of the pulse-height discriminator can be set at a level such that the smallest particle-pulse exceeds the threshold by a wide margin (because of the small inherent background in the multiplier) so that these pulses are recorded in the counting circuit. A change in the gain of the multiplier shifts the pulse-amplitude spectrum; however, the pulse counting rate does not change because all pulses are above the discriminator threshold. When a multiplier is used as a dc amplifier, however, the amplifier output signal depends on the multiplier gain.

4. APPLICATION OF ELECTRON MULTIPLIERS IN MASS SPECTROMETERS

It is possible to increase the sensitivity of mass spectrometers by approximately 3 orders of magnitude when electron multipliers are used as detectors.

Figure 7 shows a block diagram of detection apparatus used in various forms for the detection of ion beams in mass spectrometers.[14,79] The multiplier used in reference 14 has 20 stages connected in an Allen arrangement. The emitters of the multipliers are made of slightly-activated beryllium copper, thus making the gain relatively insensitive to the effect of air. The gain of the multiplier is 3.6×10^7 (mean gain of 2.6 per stage). The anode of the multiplier is also connected to a dynamic electrometer (connected in turn to an automatic recorder) and to a pulse preamplifier with a gain of 10. The main pulse amplifier has a gain of 300. The pulses are fed from the amplifer to a scaler and then to an output meter and recorder. The switching system is arranged so that the output meter reads all pulses or every tenth, hundredth, or thousandth pulse. The pulse integrator is used to record pulses at very low repetition rates, when the output meter cannot be used. Ion currents of the order of 10^{-15} amp are recorded simultaneously by the electrometer and the counting circuit. At ion currents greater than 10^{-13} amp only the electrometer is used. In the detection system described here the counting losses are 0.1% for 82,000 pulses per minute and 1% for 820,000 pulses per minute.

When an electron multiplier is used as a particle counter it is possible to detect currents between 10^{-19} amp and 10^{-14} amp and to measure isotope ratios of the order of 10^5. It is also possible to measure higher isotope ratios when isotopes are present in amounts such that comparison measurements can be made.

To measure isotope ratios in a mass spectrometer in which the ion detector is an electron multiplier, it is necessary to take account of the dependence of output signal on the mass of the ion being

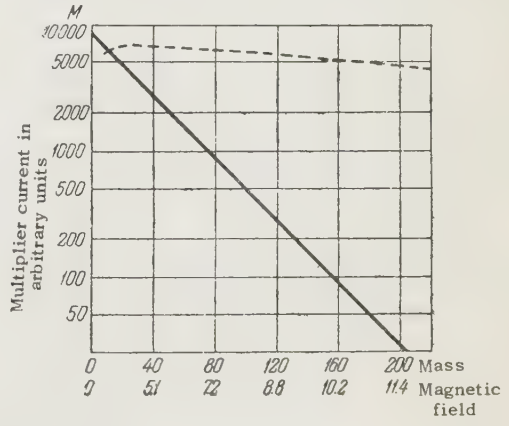


FIG. 8. Output current of a multiplier as a function of the fringing magnetic field and the mass of the detected ions (solid curve). The dashed curve is the calculated dependence of multiplier output current on ion mass when there is no magnetic field.

detected. At the present time, however, this question has still not been settled. Ploch and Walcher[13] propose that when the electron multiplier is used as a dc amplifier the following correction must be introduced for the ion mass in order to determine the true isotope ratio R_{true}:

$$R_{true} = R_{meas}\left(1 + q\frac{M_M - M_a}{M_a}\right),$$

where R_{meas} is the measured isotope ratio

$$q = \left(\frac{\partial \gamma_a}{\partial u}\right)_{M_a} : \left(\frac{u_0}{\gamma_a(u_0)}\right),$$

u is the ion accelerating voltage, u_0 is the ion accelerating voltage during the measurement, M_a and M_r are respectively the masses of the abundant and rare isotopes $\gamma_a(u_0)$ is the value of γ when the cathode is bombarded by ions of the abundant isotope accelerated by the voltage u_0. The quantity q can be determined by measuring i_0, the output current of the multiplier when the abundant isotope is detected at an accelerating voltage u_0, and $i_0 + di_0$, the current when the voltage is changed by an amount du. We find that $q = di/i_0 : du/u_0$. The quantity q depends on the state of the cathode surface and on the ion energy and varies from 1 to $\frac{1}{3}$.

The need for introducing corrections when determining isotope ratios by a mass spectrometer in which an electron multiplier is used has been indicated by a number of authors.[1,2,92] At the same time, White and Collins[14,46] and Andreeva[79] report no distortion in the isotope ratios when detecting ions with electron multipliers. The same is reported by Leland[45] and Kennett and Thode,[81,92] who used an electron-multiplier mass spectrometer to investigate the rare gases formed by fission of U^{232}, U^{235}, and Pu^{239}.

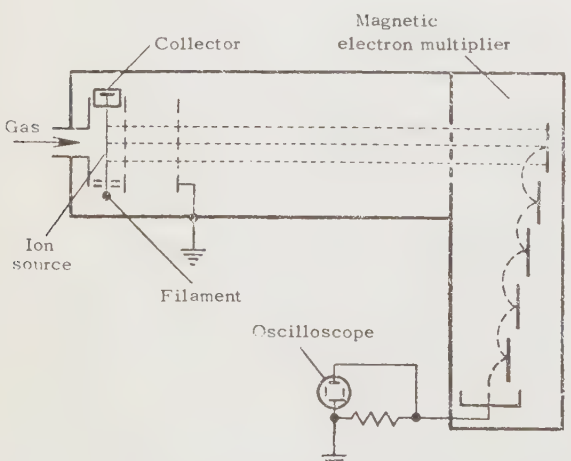


FIG. 9. Diagram of a radio-frequency mass spectrometer with a magnetic electron multiplier.

The use of an electron multiplier as a particle counter for individual ions yields a significant reduction in the difference between the measured and true isotopic ratios. When ions of the different isotopes are detected there is a change in the mean amplitude of the multiplier pulses. However, since there is a pulse-height discriminator at the output of the amplifier, and since this unit produces a standard signal regardless of the amplitude of the input pulses, the counting rate remains unchanged if the discriminator threshold is set below the level of the smallest pulses. On the other hand, when the energy of the detected ions is less than 10 kev, as is usually the case in mass spectrometers, the ion counting efficiency may be different for different isotopes, since ions of different isotopes eject different numbers of secondary electrons from the multiplier cathode.

In considering the use of a multiplier in a mass spectrometer one must keep also in mind the fact that the magnetic fringing field of the spectrometer can have a strong effect on the operation of the electron multiplier, causing defocusing of the electron beam. If there is a very strong fringing field the multiplier may be completely inoperative. Figure 8 shows curves for γ at the cathode of the multiplier (dotted curve) and the gain of the multiplier (solid curve) as functions of the mass of the ion being detected and the strength of the spectrometer fringing field.[1] The difference between the two curves is a result of the reduction in multiplier gain as the strength of the fringing field is increased. Two methods are used to compensate for this effect. In one case the electron multiplier is placed in a magnetic shield. In the second case an auxiliary magnet is placed near the multiplier in such a way as to set up a field which compensates for the fringing field of the mass spectrometer.

Special magnetic electron multipliers have been devised for operation in uniform magnetic fields up to several hundred oersteds.[43]

Because of their high resolving power electron multipliers are widely used in magnetless radio-frequency mass spectrometers; with these instruments it is possible to examine the mass spectrum of any material in a short period of time. The spectrum is observed on the screen of an oscilloscope. The multipliers in these instruments are pulsed and detect ion bunches corresponding to pulse lengths of tenths of a microsecond.[39,47-49]

Figure 9 shows a diagram of a magnetless, radio-frequency mass spectrometer in which a magnetic electron multiplier is used.[48] The cathode in this multiplier is a plane surface; with this ge-

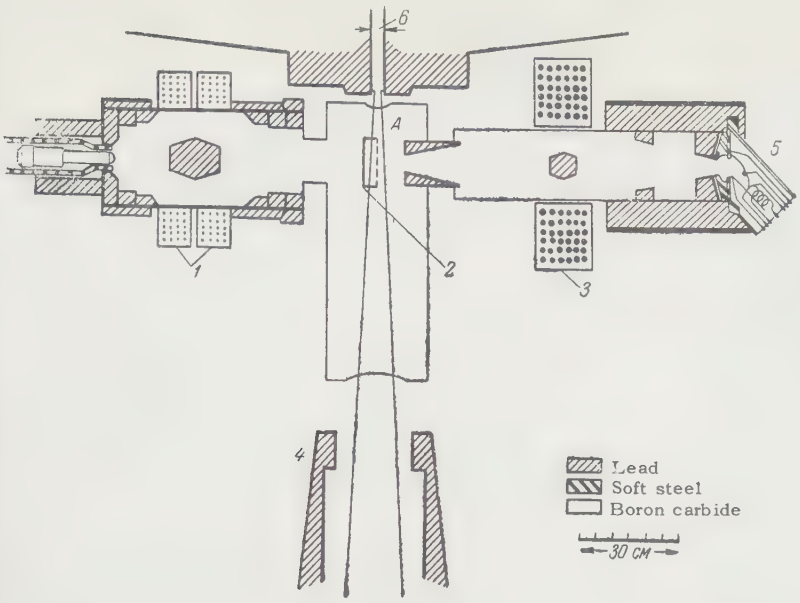


FIG. 10. Apparatus for the investigation of β decay of the neutron. 1) Electron spectrometer; 2) semi-cylindrical electrode; 3) proton spectrometer; 4) beam collector; 5) electron multiplier; 6) cross-section of the neutron beam.

ometry one can minimize effects which tend to increase the duration of the multiplier pulse Δt_i because of the non-simultaneous arrival of ions at the cathode of the multiplier. This is to be contrasted with other multiplier systems in which curved cathodes are used. Thus, for example, in a standard electron multiplier with trough-shaped electrodes we find that with a cathode potential of -4 kev and ions of mass 100, $\Delta t_i \approx 5 \times 10^{-8}$ sec.[50] In the case of an electron multiplier with "venetian blind" electrodes we find $\Delta t_i \approx 2 \times 10^{-8}$ sec under the same conditions.[50] The computed length of the pulse from a magnetic electron multiplier is 1.3×10^{-10} seconds;[43] this value is considerably smaller than that in electron multipliers with focusing electrodes. For this reason, it is possible to use in magnetic electron multipliers a large number of emitters of non-activated beryllium-copper with an average gain per stage of the order of two. The gain of such a multiplier is not affected if the electrodes are exposed to air or to various gases. The magnetic electron multipliers used in mass spectrometry have varying numbers of stages and overall gains ranging from 10^5 to 5×10^9. The pulses from multipliers with high gains are applied directly to the plates of a cathode-ray tube.

In the analysis of mass spectrograms it is necessary to take account of the fact that the output signal from the electron multiplier is different for different ions. To obtain precise values in mass analysis it is necessary to calibrate the mass spectrometer with a gas of known composition.[51]

Magnetic electron multipliers are also used for detecting ions in dynamic mass spectrometers such as the chronotron.[52] A basic disadvantage of magnetic multipliers is the fact that these devices require a uniform magnetic field of several hundred oersteds.

Electrostatic electron multipliers are also of interest. A recently developed electron multiplier in an image converter[53] is characterized by an extremely short pulse. An experimental version of this device has been constructed in the form of a demountable photo-multiplier; however this multiplier can be used for ion detection when the photocathode is replaced by an emitter which emits electrons directly when struck by charged particles. Amplification of the electron flow in the multiplier is achieved by secondary emission from the "back" of thin film emitters. The measured rise time for pulses in the multiplier is 1×10^{-9} sec; according to the design calculations this time should be of the order of 10^{-11} sec. The high resolution is obtained by means of the high interstage voltages and simple geometry.

Electron multipliers have also been used successfully for detecting ions in radio-spectroscopy devices in the determination of magnetic moments of different isotopes contained in a mixture of isotopes.[54-57]

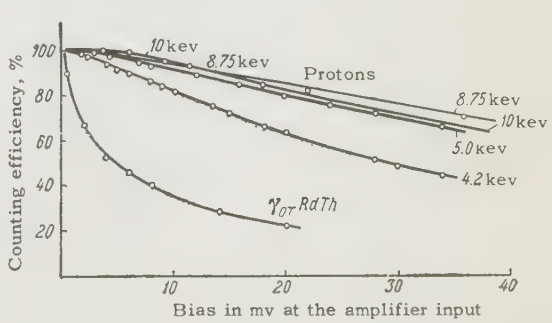


FIG. 11. Discrimator curves for the detection of protons and γ rays by an electron multiplier.

5. APPLICATION OF ELECTRON MULTIPLIERS IN NUCLEAR PHYSICS

The experimental study of different β transformations requires that one detect recoil nuclei with an initial energy of several tens or hundreds of electron volts. Electron multipliers turn out to be the most convenient detectors for such slow ions. In particular, electron multipliers have been used successfully in studying the β decay of the free neutron.[58-60] A diagram of the apparatus used in a number of investigations of β-decay of free neutrons[58] is shown in Fig. 10. The recoil protons formed in the neutron beam are focused from region A into the input window of a proton spectrometer; a positive potential of approximately 10 to 15 kv is applied to the semi-cylindrical electrode in the spectrometer. At the output of the proton spectrometer there is a beryllium-copper electron multiplier which detects the recoil protons. The electrons formed in the β decay of the neutrons are detected by a scintillation counter. The pulses from both counters are applied to a coincidence circuit. The background of this experiment is very large because of the scattered γ radiation. Figure 11 shows discriminator curves obtained by using an electron multiplier with beryllium-copper emitters; these data pertain to protons with energies ranging from 4 to 10 kev and γ rays from Ra-Th.[35] From the behavior of these curves it is apparent that wide amplitude ranges are obtained when a multiplier is used to detect protons and γ rays. Although an amplitude discriminator can be used in helping to distinguish the proton pulses from the γ background this procedure leads to a considerable reduction in the proton counting efficiency.

Electron multipliers have been used widely in detecting recoil nuclei formed in β transformations of various radioactive isotopes. The most convenient examples are the radioactive isotopes He^6, Be^7, Ne^{19}, A^{37}.[61-70] The energy of the recoil nuclei in different experiments has been measured by means of an electrostatic spectrometer or spectrometers in which measurements are made of the time of flight of the recoil proton from the source to the cathode of the electron multiplier. When electron multipliers are used in these experiments in cases in which the cathode of the electron multiplier is in line with the source of recoil nuclei, the multiplier background increases significantly when the source is heated. This is apparently due to detection of ions formed at the surface of the heated target. In another case[62] it has been noted that when an electron multiplier is used to record Li^+ recoil nuclei the multiplier background during the measurements varies from 4 to 40 pulses per minute; this is apparently due to the increase in thermal emission from the cathode of the multiplier as a consequence of the deposition of a lithium layer. According to quantitative estimates of the number of recoil nuclei incident on the cathode of the multiplier, it is necessary to make corrections because some of the recoil nuclei are intercepted by a grid which is generally used to obtain a more uniform electric field in front of the input window of the multiplier. In addition, it is necessary to take account of the fact that the different recoil nuclei are charged and that the efficiency of the counter may also depend on the charge of these nuclei.

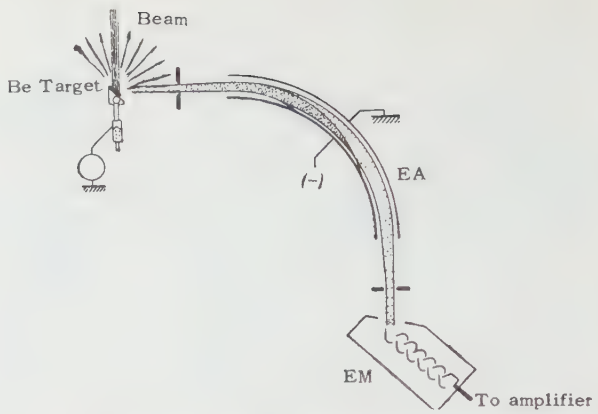


FIG. 12. Apparatus used to investigate the energy spectra of nuclear reaction products. EA) electrostatic analyzer, EM) electron multiplier.

In a number of cases electron multipliers have been used for detection of heavy charged particles which are produced in nuclear reactions caused by the bombardment of a target by a beam of accelerated ions.[71-73] Figure 12 shows a diagram of an experiment carried out by del Rossario[71] to determine the energy of products of nuclear reactions produced in the bombardment of a beryllium target by protons with energies of 239, 268, and 397 kev. The energy of the charged particles formed in the $Be^9(p,\alpha)Li^6$ and $Be^9(p,n)Be^8$ was determined by means of cylindrical electrostatic analyzer. The electron multiplier was used to detect particles which passed through the analyzer. The multiplier had 12 beryllium-copper dynodes. The energy spectra of the lithium ions formed in these nuclear reactions are shown in Fig. 13. The high counting rate in the low-energy region is due to the scattered protons; because of this background it was impossible to determine the position of the Li^{+3} peak for bombardment by 397-kev protons.

Tulinov[73] investigated the angular distribution

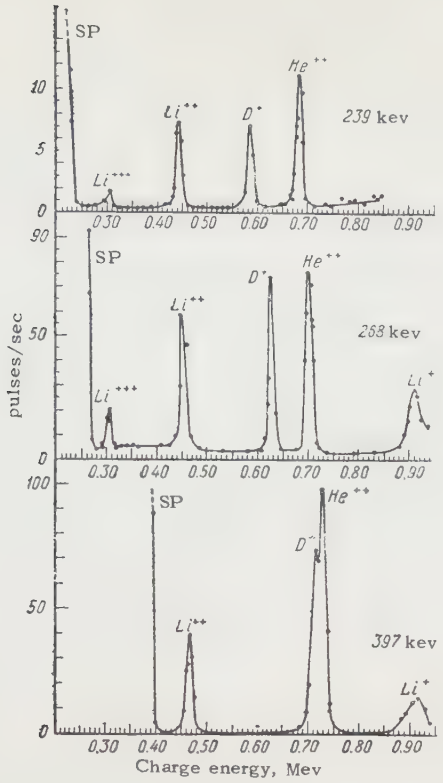


FIG. 13. Energy spectra of charged particles produced in nuclear reactions induced by protons with energies of 33, 239, 268, and 397 kev. SP) scattered proton.

of recoil nuclei from the $Be^9(d, \alpha)Li^7$ reaction; it was found possible to use the angle method to study the energy levels of the recoil nuclei. In this experiment a beam of deuterons accelerated to approximately 4 Mev in a cyclotron passes into the experimental chamber and bombards a beryllium target. The α particles formed in the nuclear reactions are detected with a proportional counter and the Li^7 recoil nuclei, with energies of approximately 1 Mev, are recorded by an electron multiplier with beryllium-copper emitters. The pulses from both counters are applied to a coincidence circuit. During the measurements it was found that the background loading of the multiplier was very large, so that it was difficult to obtain precise data. When thin organic films, through which the recoil nuclei could pass easily, are placed in front of the input window of the multiplier the background is reduced substantially.

Apparently the high background when there is no organic film in front of the input window results from the detection of slow ions or ultraviolet photons produced as the deuteron beam passes through the target and working volume of the experimental chamber.

In certain experiments it is necessary to detect slow ions against a background of intense electromagnetic or particle radiation. In these cases the desired radiation can be separated by two-channel electron multipliers connected in coincidence or anti-coincidence.[74,75] In addition to being used in nuclear physics, electron multipliers have also been used in the investigation of the effects that arise in the passage of low-energy ions through matter.[76,77] In a number of investigations multipliers have been used to detect the neutral component of an ion beam which is formed when the ions pass through metal foils[28] or a gaseous medium.[78]

Electron multipliers with alloy emitters are characterized by high sensitivity to ions. In addition, these devices are relatively insensitive to exposure to air. For this reason it is possible that they will find application in the detection of weak ion currents in ionization gauges, for example in the high-vacuum radioactive ionization gauge described in reference 80.

We have not considered other applications of electron multipliers since these applications lie beyond the scope of the present review. However, we may mention the fact that multipliers with alloy emitters have also been used in the detection of electron beams in raster electron microscopes,[86] in secondary-emission vacuum tubes,[87,88] in television transmitting tubes, in studying the emission of electrons after mechanical processing of metal surfaces,[89] and so on. In addition, these multipliers find wide application in the detection of ultraviolet and soft x-rays.[36,90,91]

[1] Inghram, Hayden, and Hess, "Mass Spectroscopy in Physics Research" National Bureau of Standards Circular No. 522 (1951) p. 257.

[2] Higatsberger, Demorest, and Nier, J. Appl. Phys. **25**, 883 (1954).

[3] Barnett, Evans, and Stier, Rev. Sci. Instr. **25**, 1112 (1954).

[4] Aarset, Cloud, and Trump, J. Appl. Phys. **25**, 1365 (1954).

[5] M. A. Eremeev and N. N. Petrov, **Проблемы современной физики** (Problems in Contemporary Physics) No. 9, 133 (1957).

[6] E. S. Mironov and L. M. Nemenov, J. Exptl. Theoret. Phys. (U.S.S.R.) **32**, 269 (1957), Soviet Phys. JETP **5**, 188 (1957).

[7] A. I. Akishin, J. Tech. Phys. (U.S.S.R.) **28**, 776 (1958), Soviet Phys. JTP **3**, 724 (1958).

[8] Hill, Buechner, Clark, and Fisk, Phys. Rev. **55**, 463 (1939).

[9] H. Geiger, Handbuch der Physik **24**, 171 (1927).

[10] L. H. Linford, Phys. Rev. **47**, 279 (1935).

[11] W. Ploch, Z. Naturforschung **5a**, 570 (1950).

[12] W. Ploch, Z. Physik **130**, 174 (1951).

[13] W. Ploch and W. Walcher, Rev. Sci. Instr. **22**, 1028 (1951).
[14] F. A. White and T. L. Collins, Appl. Spectroscopy 8, 17 (1954).
[15] Stanton, Chupka and Inghram, Rev. Sci. Instr. 27, 109 (1954).
[16] V. M. Lovtsov and A. S. Smirnov, J. Tech. Phys. (U.S.S.R.) **23**, 1737 (1953).
[17] V. M. Dukel'skii and E. Ya. Zandberg, J. Exptl. Theoret. Phys. (U.S.S.R.) **19**, 731 (1949).
[18] H. W. Berry, Phys. Rev. **74**, 848 (1948).
[19] Stier, Barnett and Evans, Phys. Rev. **96**, 973 (1954).
[20] I. P. Flaks, J. Tech. Phys. (U.S.S.R.) **25**, 2463 (1955); **25**, 2467 (1955).
[21] Massey and Burhop, Electronic and Ionic Impact Phenomena, (Russ. Transl.) London, Oxford University Press, 1958.
[22] H. D. Hagstrum, Phys. Rev. **89**, 244 (1953); **91**, 543 (1953).
[23] H. D. Hagstrum, Phys. Rev. **96**, 325 (1954); **96**, 336 (1954).
[24] J. S. Allen, Phys. Rev. **55**, 966 (1939).
[25] M. A. Eremeev and M. V. Zubchaninov, J. Exptl. Theoret. Phys. (U.S.S.R.) **12**, 359 (1942).
[26] A. H. Morrish and J. S. Allen, Phys. Rev. **74**, 1260 (1948).
[27] J. S. Allen, Proc. Inst. Radio Engrs. **38**, 346 (1950).
[28] G. A. Dissanaike, Phil. Mag. **44**, 1051 (1953).
[29] B. Kurrelmeyer and L. Hayner, Phys. Rev. **52**, 952 (1937).
[30] A. I. Akishin, Радиотехника и электроника (Radio Engineering and Electronics) **2**, No. 12, 1556 (1957).
[31] Trebukhovskii, Ergakov, and Nesterov, Приборы и техника эксперимента (Inst. and Meas. Engg.) **1**, 75 (1956).
[32] J. A. Cowan, Can. J. Phys. **32**, 101 (1954).
[33] J. S. Allen, Rev. Sci. Instr. **18**, 739 (1947).
[34] R. P. Stone, Rev. Sci. Instr. **20**, 935 (1949).
[35] J. M. Robson, Rev. Sci. Instr. **19**, 865 (1948).
[36] Morrish, Williams, and Darby, Rev. Sci. Instr. **21**, 884 (1950).
[37] T.M. Lifshitz, Радиотехника и электроника (Radio Engineering and Electronics) **1**, 1272 (1956).
[38] B. S. Kul'varskaya, Izv. Akad. Nauk SSSR, Ser. Fiz. **20**, 1029 (1956). [Columbia Tech. Transl. **20**, 931 (1956)]. Радиотехника и электроника (Radio Engineering and Electronics) **1**, 512 (1956).
[39] E. I. Agishev and N. I. Ionov. J. Tech. Phys. (U.S.S.R.) **26**, 203 (1956), Soviet Phys. JETP **1**, 201 (1956).
[40] D. T. F. Marple, Rev. Sci. Instr. **26**, 1205 (1955).
[41] M. G. Inghram, Preliminary Report No. 10, Nuclear Science Series (1955).
[42] A. I. Akishin, Приборы и техника эксперимента (Inst. and Meas. Engg.) No. 3, 72 (1957).
[43] L. G. Smith, Rev. Sci. Instr. **22**, 166 (1951).
[44] A. A. Cohen, Phys. Rev. **63**, 219 (1943).
[45] W. T. Leland, Phys. Rev. **76**, 992 (1949); **77**, 634 (1950).
[46] White, Collins, and Rourke, Phys. Rev. **97**, 566 (1955).
[47] M. M. Wolf and W. E. Stephens, Rev. Sci. Instr. **24**, 616 (1953).
[48] W. C. Wiley and I. H. McLaren, Rev. Sci. Instr. **26**, 1150 (1955).
[49] J. H. Reynolds, Rev. Sci. Instr. **27**, 928 (1956).
[50] E. I. Agishev, Dissertation, 1956.
[51] E. G. Leger, Can. J. Phys. **93**, 74 (1955).
[52] L. G. Smith, Rev. Sci. Instr. **22**, 115 (1951).
[53] E. J. Sternglass, Rev. Sci. Instr. **26**, 1202 (1955).
[54] Davis, Nagle, and Zacharias, Phys. Rev. **76**, 1068 (1949).
[55] Eisinger, Bederson, and Feld, Phys. Rev. **86**, 73 (1952).
[56] H. Lew, Phys. Rev. **91**, 619 (1953).
[57] G. Wessel and H. Lew, Phys. Rev. **92**, 641 (1953).
[58] J. M. Robson, Phys. Rev. **83**, 349 (1951).
[59] J. M. Robson, Phys. Rev. **100**, 933 (1955).
[60] Shell, Pleasonton, and McCord, Phys. Rev. **78**, 310 (1950).
[61] C. W. Sherwin, Phys. Rev. **75**, 1799 (1949); **82**, 52 (1951).
[62] R. Davis Jr., Phys. Rev. **86**, 976 (1952).
[63] G. W. Rodeback and J. S. Allen, Phys. Rev. **86**, 446 (1952).
[64] P. B. Smith and J. S. Allen, Phys. Rev. **81**, 381 (1951).
[65] J. S. Allen, Phys. Rev. **61**, 692 (1942).
[66] Allen, Paneth, and Morrish, Phys. Rev. **75**, 570 (1949).
[67] J. S. Allen and W. K. Jentschke, Phys. Rev. **89**, 902 (1953).
[68] A. H. Snell and F. Pleasonton, Phys. Rev. **97**, 246 (1955).
[69] B. M. Rustad and S. L. Ruby, Phys. Rev. **89**, 880 (1953).
[70] W. P. Alford and D. R. Hamilton, Phys. Rev. **95**, 1351 (1954).
[71] L. del Rossario, Phys. Rev. **74**, 304 (1948).
[72] H. V. Argo, Phys. Rev. **74**, 1293 (1948).
[73] A. F. Tulinov, J. Exptl. Theoret. Phys. (U.S.S.R.) **31**, 698 (1956), Soviet Phys. JETP **4**, 596 (1957).
[74] Akishin, Vasil'ev and Mikhaleva, Приборы и техника эксперимента (Inst. and Meas. Engg.) No. 5, 36 (1957).

[75] A. I. Akishin, Приборы и техника эксперимента (Inst. and Meas. Engg.) No. 5, 38 (1957).
[76] S. D. Warshaw, Phys. Rev. **79**, 1756 (1949).
[77] H. A. Wilcox, Phys. Rev. **74**, 1743 (1948).
[78] Carbone, Fuls, and Evehart, Phys. Rev. **102**, 1524 (1956).
[79] M. G. Andreeva, Приборы и техника эксперимента (Inst. and Meas. Engg.) No. 2, 53 (1958).
[80] A. I. Akishin, Приборы и техника эксперимента (Inst. and Meas. Engg.) (1958) (in press).
[81] T. J. Kennett and H. G. Thode, Phys. Rev. **103**, 323 (1956).
[82] C. F. Barnett, Phys. Rev. **91**, 224 (1953).
[83] Iu. A. Dunaev and I. P. Flaks, Dokl. Akad. Nauk SSSR, **91**, 43 (1953).
[84] G. Couchet, Compt. rend. **236**, 1862 (1953).
[85] Атомная энергия (Atomic Energy) No. 4, 168 (1956).
[86] MacMillan, Electron. Engng **25**, N 300, 46 (1956).
[87] Oberback, Wireless Engr. **28**, 331, 114 (1951).
[88] D. N. Venerovskii, J. Tech. Phys. (U.S.S.R.) **28**, 1089 (1958), Soviet Phys. JTP **3**, 1016 (1958).
[89] J. Lohff, and H. Raether, Naturwiss. **42**, 66 (1953).
[90] H. E. Hinteregger and K. Watanbe, J. Opt. Soc. Am. **43**, 604 (1953).
[91] A. M. Tyutikov and A. I. Efremov, Dokl. Akad. Nauk SSSR, **118**, 286 (1958), Soviet Phys. "Doklady" **3**, 154 (1958).
[92] G. W. Wetherill, Phys. Rev. **96**, 679 (1954).

Translated by H. Lashinsky

New Instruments and Methods of Measurement

DIFFUSION CLOUD CHAMBERS

V. K. LYAPIDEVSKII

Usp. Fiz. Nauk **66**, 111-129 (September, 1958)

CONTENTS

THE recording of ionizing particles by means of droplets condensing on ions is one of the fundamental procedures of experimental nuclear physics. The method is based on the fact that the growth of droplets to visible dimensions on charged centers occurs at smaller supersaturation than in the case of neutral particles. In the familiar Wilson cloud chamber the adiabatic expansion of a mixture of gas and vapor is employed to achieve the supersaturation required for the condensation of droplets on ions.

A number of shortcomings limit the possibilities of the Wilson cloud chamber method, the principal one being the long period of insensitivity after expansion, the complicated construction, and the susceptibility to contamination by uncharged condensation centers. The last of these factors greatly complicates the adjustment and operation of an expansion chamber. The long recovery time after expansion makes it difficult to use Wilson chambers in conjunction with particle accelerators.

The need for an instrument free of the intrinsic shortcomings of the conventional cloud chamber has led to a search for methods of achieving supersaturation other than by adiabatic expansion. A number of attempts to construct a continuously sensitive chamber have been reported. Reference 1 describes a chamber in which supersaturation results from the interdiffusion of water vapor and hydrochloric acid. This chamber did not yield satisfactory results. Reference 2 describes the so-called diffusion cloud chamber in which supersaturation is achieved by the diffusion of vapor through a condensing gas from a heated horizontal surface to a cooled surface.

The first satisfactory photographs of particle tracks in a diffusion cloud chamber were obtained in 1951 and 1952. Since then an increasing number of papers have appeared on work done by means of diffusion chambers as well as on investigations of the physical processes that occur in such chambers. The increased interest in diffusion chambers results not only from the fact that they are largely free of the shortcomings of existing Wilson cloud chambers, but also because some of the properties of diffusion chambers permit a considerably broadened use for this method of particle registration.

1. PHYSICAL PROCESSES IN DIFFUSION CLOUD CHAMBERS

A diffusion cloud chamber is a closed vessel filled with a mixture of vapor and gas (Fig. 1). Vapor sources are located close to the upper and

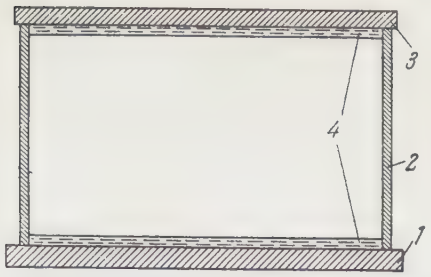


FIG. 1. Diffusion chamber: 1) Bottom; 2) Walls; 3) Cover; 4) Working liquid.

lower surfaces. When the temperature T_1 of the cover is unequal to the temperature T_2 of the bottom the following processes occur within the chamber:

1) Non-isothermal vapor diffusion through the gas from the hot to the cold surface.

2) Heat transport through the gas and vapor mixture.

3) The formation and diffusion of neutral and charged condensation nuclei.

4) Vapor condensation and droplet growth on neutral and charged nuclei.

5) Droplet motion in the gravitational field.

6) Convection (in some cases).

Amelin[3] has suggested that the temperature and partial-pressure fields in the chamber can be represented by one-dimensional equations of isothermal diffusion and thermal conduction, neglecting other processes. With these assumptions the temperature and partial pressure distributions are linear.

A better approximation was obtained by Langsdorf,[4] who took into account the temperature dependence of diffusion and thermal conductivity. Figures 2 and 3 show the temperature distribution and partial-pressure curves for two different vapor fluxes, assuming no condensation. Greater vapor flux increases both the nonlinearity of the temperature distribution and the partial pressure in the chamber. Figure 4 shows the corresponding supersaturation curves. The dot-dash curves represent the supersaturation S_1 at which droplets begin to grow on ions, and the supersaturation S_2, at which a large number of droplets are formed on uncharged nuclei.

The figure shows that supersaturation near the bottom of the chamber exceeds S_1 and S_2. The boundary of the sensitive layer is determined by the intersection of curves S_x and S_1 (points A and B on Curve 2). With greater vapor flux the partial pressure, and accordingly the supersaturation, increases (Curve 1). It is evident that as a result of condensation on charged and uncharged nuclei the true supersaturation in a diffusion chamber will be considerably smaller than is shown in Fig. 4.

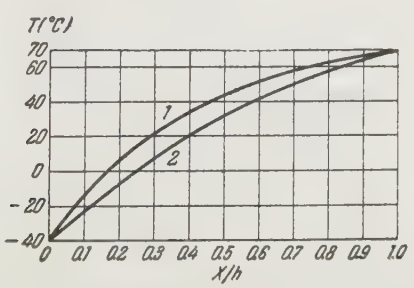


FIG. 2. Calculated temperature distributions in the volume of a chamber filled with CO_2 and CH_3OH, without condensation. Curve 1) Vapor flux 6.86×10^{-6} g/cm² sec; Curve 2) Vapor flux 2.56×10^{-6} g/cm² sec.

a) Effect of Condensation

Condensation within the chamber results in reduced partial pressure. Moreover, the release of a large amount of heat through condensation can produce a significant change in the temperature distribution.

Succi and Tagliaferri[5] have developed Langsdorf's theory, using successive approximations to determine the effect of condensation on the temperature distribution and supersaturation distribution. Figure 5 shows the temperature distribution curves for different condensation conditions. The temperature distribution is seen to change with increasing number of condensation nuclei, the temperature rising at all distances from the bottom. The temperature increases by 15 to 20°C when the ion load is large. Argan et al have further developed Langsdorf's theory, but their results do not differ qualitatively from those of earlier investigators.

References 7, 8, and 9 describe experimental studies of the effect of condensation on the temperature distribution. The temperature distributions were compared in an operating chamber and in the same chamber without condensation (the vapor source being absent). A chamber con-

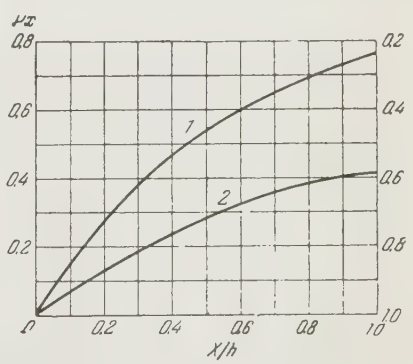


FIG. 3. Calculated partial pressure distributions, under the same conditions as for Fig. 2.

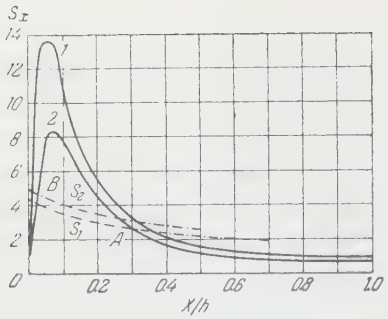


FIG. 4. Supersaturation distribution in the chamber. Curve S_2) limit of fog formation; Curve S_1) limit of condensation on ions; A and B) upper and lower boundaries of the sensitive layer.

taining rigidly fastened thermocouples was filled with air and other gases; the total pressure, vapor flux, and number of charged condensation nuclei could be varied within broad limits. It was established that when the product of the Grashof number (Gr) and Prandtl number (Pr) is at least 10^6 condensation has a negligibly small effect on the temperature distribution. When Gr Pr $\ll 10^6$ the effect of condensation on temperature distribution becomes appreciable. Figure 6 shows that when condensation occurs in a low-pressure chamber the temperature can change by as much as 20°C.

The foregoing results make it convenient to divide diffusion chambers into two types, those in which condensation has practically no effect on the temperature distribution and those in which the effect is great. The first type includes all high-pressure chambers, i.e., the great majority of those in use at the present time. Chambers with very low gas pressure are the second type. In the first type convective heat exchange with the walls plays the dominant part in the establishment of the temperature distribution. In the second type (hereinafter called low-pressure chambers) convective heat exchange with the walls plays a minor part. Langsdorf's theory[4] and similar theories,[5,6] which neglect convection, are not applicable to the

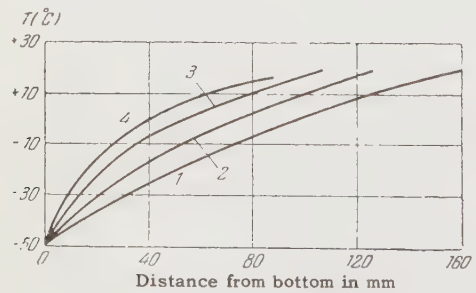


FIG. 5. Effect of condensation on temperature distribution. Curve 1) temperature distribution in the absence of condensation. Curves 2, 3, 4) temperature distribution when the number of ions passing through 1 cm² of cross section per second is 20, 60, and 180, respectively.

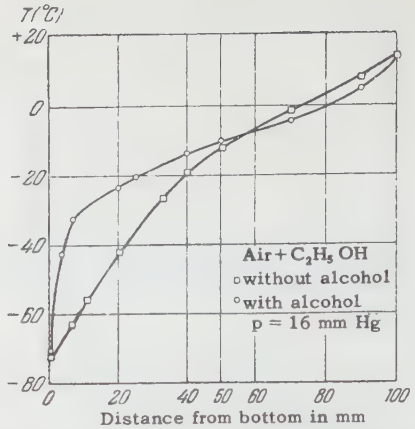


FIG. 6. Effect of condensation on temperature distribution.

ordinary diffusion chamber but can evidently be employed in analyzing processes within low-pressure chambers.

Our further discussion will concern ordinary diffusion chambers. Certain characteristics of low-pressure chambers will be considered separately.

b) Temperature Distribution and Partial Pressure Distribution

There have been several investigations of the temperature distribution in diffusion chambers[7,8,9,10]. In references 7 and 8 it was definitely shown that the temperature distribution in the space of a chamber practically coincides with the temperature distribution on the walls and can thus be regulated within broad limits. Figure 7 shows the temperature distribution curves within a chamber filled with air and ethyl alcohol vapor with a

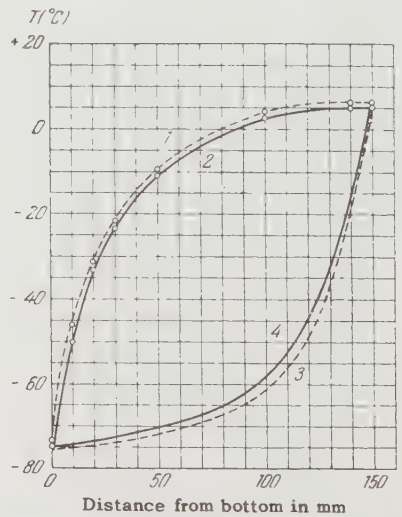


FIG. 7. Temperature distribution along walls (dashed Curves 1 and 3) and in chamber volume (Curves 2 and 4). Curves 1 and 2 were obtained in a chamber with glass walls at room temperature externally; Curves 3 and 4 were obtained in the same chamber with cooled walls.

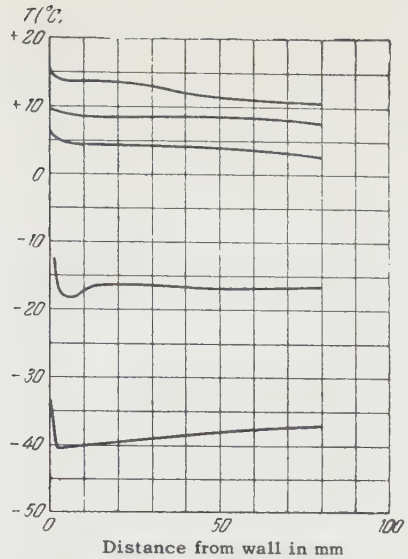


FIG. 8. Horizontal temperature distribution in a chamber that is in contact with air at room temperature.

total pressure equal to one atmosphere. Curves 1 and 2 were obtained with the outside walls of the chamber at room temperature. The temperature distribution within the chamber is seen to differ very little from that on the walls, the wall temperature at each horizontal cross section being a little higher than the temperature at the center. The temperature distribution changes sharply with outside cooling of the walls (Curves 3 and 4). In this case the wall temperature in each horizontal cross section is a little below that at the center. Figures 8 and 9 show the horizontal temperature distribution in a chamber surrounded by air at room temperature and by cold air, respectively.

When the walls are heated to room temperature externally, the gas within the chamber moves upwards. When the walls are cooled externally the direction of gas flow is reversed. There must evi-

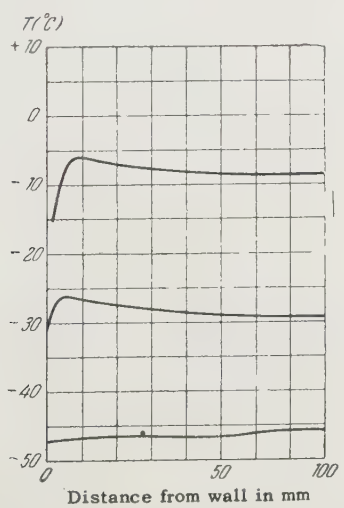


FIG. 9. Horizontal temperature distribution in a chamber with cooled walls.

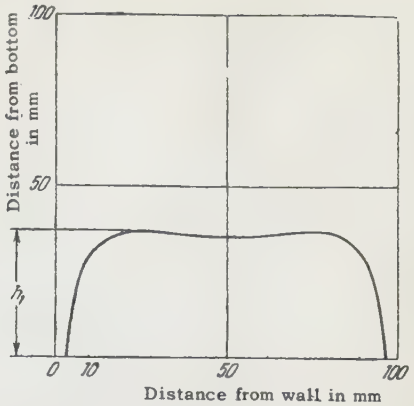


FIG. 10. Vertical section of a fog region formed after expansion.

dently be an intermediate state when the temperature distribution on the walls coincides with the temperature distribution in the gas so that convection stops. However, it is easily seen that this will not be a stable state. A change in the number of condensation nuclei within the chamber, such as following the passage of an ionizing particle, will immediately result in the liberation of heat followed by convection. Therefore convection occurs as a rule in any real chamber.

With any change of the wall-temperature distribution or with a change of the condensation conditions in the chamber the speed and in some instances the direction of convection currents will be changed. The experimental investigation described in reference 8 showed that convection within the chamber is axially symmetrical and does not lead to inhomogeneous supersaturation in the sensitive zone when the gas moves upward along the walls. The speed of the gas moving in the working volume of the chamber is comparable with the speed of falling droplet. Therefore gas motion does not appreciably distort particle tracks during the time required for their recording. Thus the temperature distribution within the chamber can be regulated without disturbing its operation.

Convective gas motion within the chamber has an important effect on the partial pressure distribution. The partial-pressure field within an operating chamber has been investigated in two ways, through the absorption of ultraviolet radiation by a gas-and-vapor mixture[11] and by the expansion method.[12] In the first method the optical density of the gas-and-vapor mixture is measured at different heights in the chamber. In the second method the degree of expansion is determined at which a fog is formed in a given region of the chamber, in which case supersaturation exceeds S_2. The upper boundary of the fog is usually nearly parallel to the bottom of the chamber. Figure 10

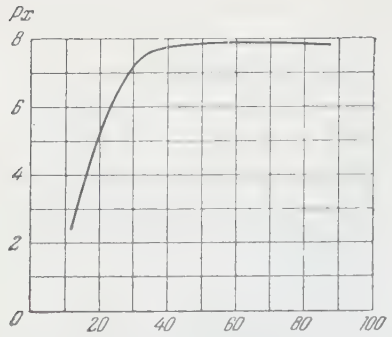


FIG. 11. Partial pressure distribution in a chamber with nonlinear temperature distribution.

shows a vertical section of a fog region formed after expansion. There is lower supersaturation in the region near the walls. The construction of the vapor source has some effect on the partial pressure distribution in the upper part of the chamber. From the distance h_1 between the upper boundary of the fog and the chamber bottom, the temperature T_{h_1} after expansion, and the supersaturation S_2 for a given gas and vapor mixture we can determine the partial pressure P_{h_1} at this boundary:

$$P_{h_1} = S_2 P_{\text{sat}}(T_{h_1}).$$

The partial pressure throughout the chamber can thus be determined by varying the degree of expansion.

Figure 11 shows a typical partial pressure distribution in a chamber with nonlinear temperature distribution. Above the condensation zone the partial pressure varies very little with height because of the strong mixing of gas and vapor.

When the temperature distribution is changed without changing the directions of convection currents the character of the partial pressure is not affected and the partial pressure above the condensation zone also remains unchanged. However, variation of the temperature distribution is accompanied by a variation of supersaturation in accordance with the change of saturated vapor pressure. This in turn affects the size of the condensation zone. For example, when the temperature distribution is varied as shown in Fig. 12, the height of the sensitive layer changes from 1.8 cm (Curve 1) to 5.5 cm (Curve 3).

It must be mentioned that a self-regulating supersaturation mechanism permits a diffusion chamber to operate in a very broad range of temperature distributions. When the temperature distribution is such that supersaturation within the chamber at any given moment is considerably above S_2 (as in Curve 1 of Fig. 4), spontaneous condensation of vapor on uncharged nuclei begins in the sensitive zone. Supersaturation decreases, reducing the number of active condensation nuclei and reducing the rate of droplet growth. The reduced consumption of vapor increases the supersaturation. As a result of this mechanism supersaturation in the sensitive zone cannot be appreciably above or below S_2 for a long time. Variation of the temperature distribution thus basically affects the magnitude of the droplet background.

In some cases the self-regulating process has a pronounced pulsating character. There are periodic variations of fog background density in the sensitive zone. This effect can be considerably reduced or entirely eliminated by a suitable change of the temperature distribution.

The operation of the chamber can also be regulated by varying the size of the vapor source through increase of the source surface or its temperature. (The vapor flux can be determined, for example, by measuring the quantity of liquid evaporated in a given time interval.) With increasing flow the partial vapor pressure must increase at each height. In the sensitive zone supersaturation cannot appreciably exceed S_2. Therefore the excess vapor is converted into the liquid phase. Thus increased vapor flow results in growth of the droplet background and has little influence on the degree of supersaturation in the sensitive zone.

Variation of the temperature distribution and vapor flow also affects the basic operating characteristics of the chamber—the maximum ion load with which the chamber can function continuously and the recovery time after a radiation pulse.

2. ION LOAD AND RECOVERY TIME

Unlike the conventional cloud chamber, the diffusion chamber is continuously sensitive. When the number of ions generated per second in one cm^3 of the chamber (ion load) does not exceed a certain limit the chamber can record all events

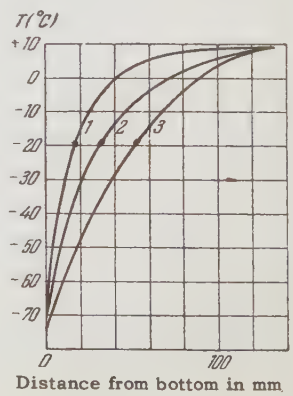


FIG. 12. Temperature distribution within chamber. The asterisk on each curve indicates the temperature at the upper boundary of the sensitive zone.

accompanied by ionization of the gas in its volume. A count of the particle tracks in a given time interval can thus be used to determine weak radioactivity in materials. Each track falls to the bottom in a few seconds, depending on the point of formation and droplet size.

Because of the continuous sensitivity and long persistence of the tracks, photographs obtained with diffusion chambers always show a certain number of background tracks. Background tracks impede analysis of the photographs and also reduce supersaturation, which for large ion loads can fall below S_1. The chamber must therefore be shielded from external radiation and the working mixture in the chamber must not be radioactive.

With small temperature gradients in the sensitive zone the chamber ceases to function at a radiation intensity comparable with the natural background. Figure 12 shows that a reduced temperature gradient is associated with increased height h of the sensitive zone. With an increase of h the number of condensation nuclei in the lower part of the chamber increases, owing to droplets falling from above. Near the bottom supersaturation decreases, resulting in the formation of an insensitive zone. Therefore the thickness of the sensitive zone in a diffusion chamber does not usually exceed 6 or 7 cm. With increasing thickness of the sensitive zone and increasing temperature gradient, the ion-load limit is raised.

Vapor flow also has an appreciable affect on the permissible ion load. Thus in a chamber filled with air and ethyl-alcohol vapor at atmospheric pressure, an increase of vapor flow by a factor of 6 increases the ion-load limit by a factor of 3. With a higher ion load the amount of vapor condensing on uncharged nuclei is reduced, thus reducing the droplet background. It is evident that the total amount of liquid condensing per unit time within the chamber remains practically constant.

Under certain operating conditions the diffusion chamber can record continuously radiation with a background that exceeds the cosmic radiation background by a factor of hundreds.[13] Under such conditions an important role is played by convection currents that transport the gas and vapor mixture from the heated top to the cold bottom.

After the passage of a charged particle through the sensitive zone, the latter contains a region with supersaturation below S_1 in the shape of a cylinder whose diameter depends on the ionizing power of the particle and the quantity of vapor at the given height. The recovery time of supersaturation after passage of a single particle is a few seconds. When many particles are involved, some of them traverse regions in which supersaturation had not received; they thus produce discontinuous tracks. When a beam consisting of very many particles passes through the sensitive zone, supersaturation can fall considerably below S_1, with the recovery time rising to 10 or 15 sec. A larger number of ions produced simultaneously in the sensitive zone is accompanied by a reduction of supersaturation and a longer recovery time. The recovery time also increases with the thickness of the sensitive zone. The recovery time can be shortened by increasing vapor flow.[14]

When the entire volume of a chamber is irradiated, condensation nuclei appear above as well as within the sensitive zone. By moving about in the chamber as far as the upper boundary of the sensitive zone, these nuclei form a more or less uniform droplet background. When the ion concentration is large, supersaturation in the sensitive zone can fall below S_1. The recovery time after cessation of irradiation depends on the nature of the vapor. This apparently indicates that the condensation nuclei are not ions but charged droplets diffusing into the sensitive zone from the region above.

Charged condensation nuclei can be swept from the space above the sensitive zone by an electric field applied between the cover serving as one electrode and a second electrode close to the upper boundary of the sensitive zone.

The chamber is cleared of uncharged condensation nuclei by migration of these nuclei into the sensitive zone, after which the drops that are formed fall to the bottom. When the number of nuclei is very large the rate of sweeping is limited by the vapor flux. When the number is small and the rate at which vapor is consumed in condensation is less than the vapor flux, the recovery time is determined by the migration rate of nuclei into the sensitive zone. Therefore recovery time depends strongly on the character of gaseous motion in the chamber and on its design. A simple chamber with a strongly nonlinear temperature distribution possesses a recovery time of the order of two minutes when filled with unpurified outside air. When the temperature distribution curve is rectified, convection velocity is reduced and the recovery time increases, reaching 20 or more minutes in some instances.

It is important to note that when a chamber contains no constant source of condensation nuclei it is more or less rapidly self-clearing. This ability clearly distinguishes the operation of a diffusion chamber and is one of its principal advantages over a conventional cloud chamber.

3. DIRECTION OF DIFFUSION

In a diffusion chamber the density varies with height and is a function of temperature and of the relative amounts of gas and vapor. When the molecular weight of the gas is smaller than that of the vapor this can result in a mixture whose density decreases in the downward direction, resulting in a singular type of convection. A system of vertical convection currents is established between the top and bottom. This occurs principally in low-pressure chambers, where the partial vapor pressure is commensurate with the total pressure.

Convection can be obviated by reversing the direction of diffusion. Chambers with a hot bottom and cold top (upward diffusion) have been described in references 15, 16, and 17. Investigations of the operation of these chambers are described in references 8 and 18.

At high gas pressures, when the vapor pressure is negligibly small compared with the total pressure, the variation of the relative composition of the mixture is small and the relation between density and height depends mainly on the temperature factor. Increase of the gas pressure or reduction of the partial vapor pressure aids stabilization of the gas in a downward-diffusion chamber. The same effect is achieved by using a gas of greater molecular weight or a vapor of smaller molecular weight. Chambers filled with hydrogen or helium are found to operate in a stable manner only when the total pressure of the mixture exceeds a few atmospheres[19] (the vapor source temperature being +10°C or +15°C). When the temperature of the top plate is reduced the partial vapor pressure is sharply lowered; this permits operation with light gases below atmospheric pressure.[20] In this last case the bottom temperature must be lowered in order to maintain the required temperature distribution of the chamber.

4. SIMILAR CHAMBERS

In high-pressure chambers, which are most widely used, the temperature distribution can be regulated over a wide range, as already mentioned, thus regulating the degree of supersaturation. Results obtained with any one chamber can be applied to another chamber of similar type. For calibration, a chamber is usually filled with air at atmospheric pressure and the optimum temperature distribution is established experimentally. This then gives optimum conditions for other gases and pressures (if Gr Pr > 10^6 in all cases). For chambers with the same temperature distribution the ion-load limit and recovery time will depend on the nature of the gas, the pressure, and vapor flux.

Shutt has established that the ion load and pressure are related to the gas parameters by the following equation:[21]

$$B = \nu_0 D_0^{-\frac{1}{3}} p^{\frac{1}{3}} (n_0 a Z p)^{\frac{4}{3}},$$

where B is a constant for a given temperature distribution, μ_0 is the kinematic viscosity at 273°C and normal pressure, D_0 is the diffusion coefficient of the vapor in the gas at normal pressure and 273°C, n_0 is the number of ions generated in one cm^3 of air per second under one atmosphere of pressure, a is the number of atoms per gas molecule, and Z is the atomic number of the gas.

Shutt's theory does not take neutral condensation nuclei into account. By making a suitable correction Bevan obtained a formula which takes both neutral and charged condensation nuclei into account.[22]

$$B_\alpha = \nu_0 D_0^{-\frac{1}{3}} p^{\frac{1}{3}} [n_0 a Z p + 14.5 e^{0.116 t}]^{\frac{4}{3}},$$

where t is the highest temperature within the chamber in degrees centigrade (usually the temperature of the vapor source).

Experimental tests of the foregoing relations have been described in references 8, 9, and 14. It was found that when one gas is replaced by another and the pressure is varied the relation between μ_0, D_0, a, Z, the maximum ion load, and the pressure is well described by Shutt's equation.

Shutt's theory does not allow for the effect of vapor flow on the operation of the chamber; in references 8 and 14 this effect was shown to be significant. The variation of vapor flux (by increasing the surface of the vapor source) can multiply the ion load limit several times and greatly reduce the recovery time.

Bevan's equation does not allow for the fact that the flow of vapor is enhanced when the vapor-source temperature is increased at the same time as the number of neutral nuclei is increased. Therefore Bevan's equation has not been confirmed experimentally. For example, with an increase of source temperature the ion-load limit in a number of instances increases[8] instead of decreasing as would follow from Bevan's equation.

According to Shutt[21] a minimum temperature gradient exists for operation of the chamber under given conditions. Bevan experimentally determined the relation between the minimum temperature gradient G in the sensitive zone and the parameter B.[22] However, his results appear to

have limited application since B depends not only on the temperature gradient in the sensitive zone but also on the temperature distribution within the entire chamber volume. For chambers of different design with the same temperature gradient in the sensitive zone, the temperature distribution within the chamber volume can differ appreciably. This evidently accounts for the fact that measurements performed on chambers of different design give a relation between G and B which differs from that of Bevan.[8]

5. LOW-PRESSURE CHAMBERS

Low-pressure chambers are described in references 20 and 23. The temperature distribution in in the gas of a low-pressure chamber exhibits very little dependence on the temperature distribution along the walls[9] and thus cannot be regulated as in high-pressure chambers. When pressure is reduced the temperature in the lower part of an operating chamber can either increase or decrease, depending on the character of the temperature distribution at the higher pressure. Pressure reduction in a chamber with a steep non-linear temperature distribution (and thus with a thin sensitive layer) reduces the temperature in the volume. On the other hand, in a chamber with a nearly linear temperature distribution pressure reduction raises the temperature in the condensation zone. In both cases at reduced gas pressure the sensitive zone is observed to be of lesser height. This effect is evidently associated with reduced partial pressure in the chamber volume.[9]

In chambers with a thick sensitive zone (> 5 cm) reduction of the zone at lowered pressure results to a considerable extent from a higher temperature of the chamber volume. It is shown in reference 23 that at lower pressure the thickness of the sensitive zone is reduced from 100 mm at atmospheric pressure to 20 mm at 30 mm Hg.

At pressures below 30 mm Hg periodic fluctuations of the background density occur.[24] The sensitive zone breaks up into cells. Fig. 13 is a typical photograph of the sensitive layer. The fog alternately clears and is restored (pulsations of the fog background) and the shapes and arrangement of the cells can vary. Cell size and pulsation frequency depend on chamber pressure and on the nature of the gas and vapor. Pulsation frequency increases with reduced vapor flow and reduced gas pressure. The smaller the latent heat of condensation and the smaller the molecular weight of the gas, the higher the pulsation frequency. Fluctuations of the fog background density are evidently associated with the liberation of heat through condensation.

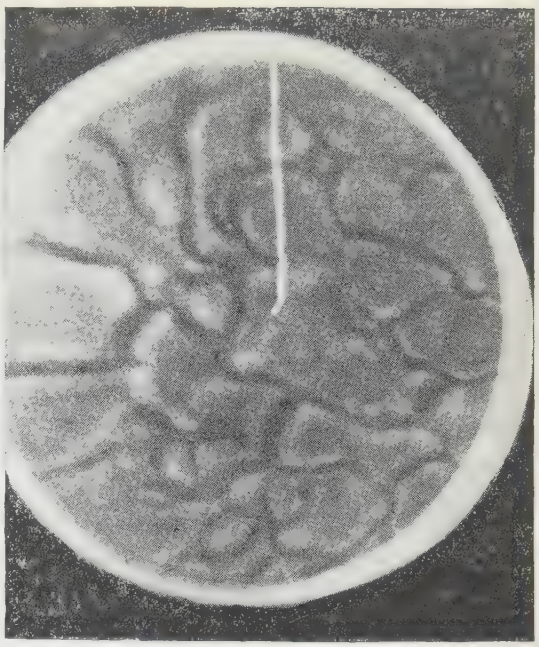


FIG. 13. Photograph of a sensitive zone in a low-pressure chamber (p = 40 mm Hg) filled with air and ethyl-alcohol vapor.

It is important to note that the track of an alpha particle passing through a few cells is continuous. Under certain conditions of illumination, a photograph may not show the cells at the same time that the alpha-particle track is clearly visible.

As already mentioned, fluctuations of the droplet background also occur in high-pressure chambers for a given temperature distribution. But unlike the case of high-pressure chambers, where this effect can be avoided, time fluctuations of the droplet background density cannot be eliminated in a low-pressure chamber since the temperature distribution in the chamber cannot be regulated.

6. CHAMBER CONTROL

A diffusion chamber can be controlled by either external or internal 25 counters. In work with accelerators the chamber is usually controlled by the accelerator itself.[26] The sequence of operations and typical time intervals between them are shown in Fig. 4, which was taken from reference 26.

In a diffusion chamber droplets begin to grow on ions immediately following the passage of a particle, whereas in a conventional cloud chamber some time elapses between the passage of the particle and the production of the required supersaturation; during this time ions diffuse from the places where they were generated. As a result particle tracks in a diffusion chamber are clearer than in

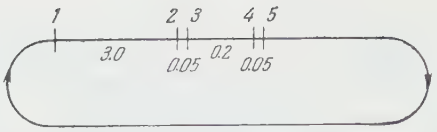


FIG. 14. Operating cycle of a diffusion chamber. Times are given in seconds. 1) Switching off of electric sweeping field; 2, 3) Entrance of particles; 4) Illumination pulse; 5) Switching on of electric sweeping field and film-moving mechanism. Total duration of cycle is 6 to 12 sec.

a conventional cloud chamber controlled by counters. It must also be remembered that in a diffusion chamber the tracks are not distorted by the motion of the expanding gas. The control system of a diffusion chamber is considerably simpler than that of a chamber involving expansion.

Because of its continuous operation, a diffusion chamber can be controlled by means of a photomultiplier "scanning" the working volume. Light scattered by the particle track enters the photomultiplier, the output pulse of which then triggers the control system. This control is subject to the condition that light scattered by the droplet background is considerably weaker than the light pulse from the particle track. This condition is easily fulfilled in the recording of strongly ionizing particles such as alpha particles or of showers generated by high-energy particles.

Reference 27 reports briefly on the possibility of controlling a diffusion chamber by means of a photocell. Control of the chamber by means of a photomultiplier is very promising, for example, for the recording of extensive air showers or events accompanied by the appearance of strongly ionizing particles in the volume of the diffusion chamber itself.

7. DESIGN OF DIFFUSION CLOUD CHAMBERS

a) Cooling of Chamber Bottom

The bottom temperature required for satisfactory operation of the chamber depends to a considerable extent on the nature of the vapor. It is customary to use methyl or ethyl alcohol vapor, for which the bottom temperature must be below −40°C. The bottom can be cooled most simply by direct contact with dry ice pressed against the bottom by a spring. This method gives very satisfactory results when the chamber walls are poor conductors. In metal chambers there is large heat flow along the walls. For example, in a chamber with a steel wall 6 mm thick and 430 mm in diameter and with a temperature gradient of 6°C per centimeter, the required heat flow is 50 cal/sec.[28] The evaporation rate of the dry ice then reaches 1 l/sec, which leads to the formation of a heat-insulating layer of gas. This in turn causes nonuniform cooling of the bottom.

Somewhat better results can be obtained by cooling the bottom with a mixture of dry ice and acetone or alcohol.[29,14] A container welded to the bottom is filled with the cooling mixture, which is in contact with the undersurface of the bottom plate, thus insuring more uniform cooling. In this case, however, the liberation of a large amount of gas with large heat flow reduces the total surface of contact with the cooling mixture, thus preventing the achievement of low temperatures.

Much lower temperatures are achieved by the use of liquid nitrogen. The bottom is then cooled either by means of a heat conductor within a Dewar vessel[20] or by passage of the liquid nitrogen through a coil of pipe fastened to the bottom.[8,12]

In diffusion chambers that are operated in conjunction with accelerators the cooling system usually consists of a heat exhanger (or refrigerator) and coil fastened directly to the bottom of the chamber. The cooling liquid is pumped through the heat exchanger, which is usually a tank filled with a mixture of dry ice and acetone, and then through the coil fastened to the bottom. The temperature can be regulated by varying the vapor pressure above the mixture; in this way the bottom can be cooled to about −100°C.[28] Reference 28 gives calculations which can be useful for the designing of a chamber-cooling circulating system. The cooling of a chamber with a transparent bottom, such as is described in reference 30, presents certain specific difficulties.

b) Chamber Walls

The walls of a diffusion chamber serve two different purposes. They form a portion of the container, which when the chamber is filled with gases at different pressures must be gastight and sufficiently strong. Secondly, the walls provide the boundary temperatures, which strongly affect the temperature distribution within the chamber.

To prevent local convection currents, which would disturb normal operation, the wall temperature in each horizontal cross section must be constant. This condition is easily fulfilled in the case of transparent walls, where no window is required for illumination. Metal walls require windows for illumination and for the admission of particles. In this case the flow of heat along the wall from the top to the bottom encounters different resistance in different cross sections. The vertical temperature distribution along the wall there-

fore varies and the existence of a horizontal component of the temperature gradient results in convection currents, which are easily detected through the formation of fog curtains that remain motionless for a long period. Convection establishes an internal temperature distribution in the chamber volume which is different from that along the walls.

Temperature differences in any horizontal cross section can be obviated through the use of sources of cold and heat sources disposed on the outside of the walls, but it is considerably simpler to create a temperature field by placing a special container inside the chamber.[8, 14, 31] A plexiglas cylinder with a pure copper bottom is placed inside the metal diffusion chamber. The bottom temperature of this cylinder is that of the copper disk, while its upper temperature is that of a copper trough containing evaporating liquid. The outside of the cylinder is surrounded by gas, which transfers heat between the different elements in the chamber. Measurements with a thermocouple sliding along the cylinder wall[8] have shown that the horizontal component of the temperature gradient does not exceed 1 or 2°C.

Difficulties associated with the need for equalizing the wall temperature of a metal chamber containing windows can be overcome to a considerable extent if illumination is achieved through the use of a window in the top cover and mirrors inside the chamber.[8] There are then no windows in the side walls. Side windows are also undesirable from considerations of strength. Large temperature gradients cause internal stresses in connected parts made of materials with different expansion coefficients. This is especially dangerous in high-pressure chambers, where leaks can occur as a result of repeated strains during cooling and heating. Methods and devices used to prevent leaks in window chambers complicate the operation of the entire instrument.[14]

c) Vapor Source

The vapor source in downward-diffusion chambers is usually a trough, fastened to the top plate, containing the working liquid. Near the top the vapor pressure decreases with increasing distance from the vapor source. To produce uniform supersaturation equal to unity near the top plate, the vapor source can be a liquid maintained at constant temperature[8] and evaporating from the lower surface of the lid. The cover then takes the form of a vessel with gastight walls and a porous bottom. A simpler vaporizer was described in reference 17. A very damp cloth was fastened to the lower surface of the metal lid, the moisture being maintained at a constant level by means of tubes.

Vapor sources in the form of troughs possess two serious disadvantages. They require periodic replenishing of the working liquid, which in the case of high or low gas pressures necessitates special devices,[14] thus complicating the design and operation of the chamber. Secondly, when a mixture of different liquids is used as a vapor source the composition of the mixture changes during operation, thus varying the operating conditions of the chamber. Consequently pure alcohol is usually used in diffusion chambers, although it is known that mixtures give better results. Unconnected troughs containing different pure liquids can be used to provide a vapor mixture, but this intensifies the disadvantage resulting from the fact that special devices are needed to replenish the working liquids.

A diffusion chamber with a continuously operating vapor source[32] and without the aforementioned shortcomings was proposed in 1952. A liquid poured on the chamber bottom rises through capillary action in porous plates located close to the walls; it then evaporates and diffuses into the chamber. This process is continuous so long as the required temperature distribution is maintained between the top and bottom. The composition of the mixture remains constant. Investigation of the structure of the sensitive zone by means of a gas discharge[8] and by means of expansion[12] in a chamber with a continuously operating vapor source[33] shows that supersaturation in the sensitive zone is practically independent of the horizontal coordinate. Partial pressure distribution within such a chamber is of the same character as in a chamber with the usual vapor source in a trough. Windows cut through the porous plates for illumination of the chamber do not cause inhomogeneity of the sensitive zone.

In a chamber with a continuous vapor source the droplet background is usually less dense and more uniform than in a chamber with a trough source close to the lid. This apparently results from the fact that under certain conditions a large number of condensation nuclei can be formed close to a vapor source provided by a trough fastened to the lid.

In chambers with a trough the vapor source temperature is usually a few degrees above the temperature of the ambient gas. This temperature difference is especially great in chambers where auxiliary heating of the vapor source is provided; see reference 22, for example. When gas from

below comes into contact with such a source the condensation nuclei that are created migrate into the sensitive zone forming a fog background.

With nonuniform heating of the lid or walls, local convection currents arise and create a corresponding configuration in the droplet background. Downward currents of cooled gas enter the sensitive zone and form fog curtain that remain almost unchanged in shape and position for long periods. Such fog curtains have been noted by many investigators.[28] When the vapor source has a lower temperature than the gas near the lid (as with a continuously operating source) this effect is much less pronounced.

A continuous vapor source can be used in large chambers; for example, a chamber with a 60 × 90 cm bottom was demonstrated at the All-Union Industrial Exposition of 1956.

It has already been noted that the number of neutral centers and thus the droplet background will grow with rising vapor source temperature. At an interior temperature of 60°C a chamber cannot furnish good tracks of ionizing particles, and at about 80°C all registration ceases.

d) Illumination of Chamber

A diffusion chamber is usually illuminated from the side through transparent walls or through windows in the walls of a metal chamber. A heat-insulating air gap is used to prevent frosting or fogging of the outside surfaces of windows.[34, 35] The temperature of the surface exposed to the outside air is then usually above the dew point. The outside walls are sometimes heated to prevent condensation. To prevent fogging of the outside surface of windows in metal chambers special jackets are usually used to insulate the windows from the outside air.[14] The illuminating lamp and lens system are mounted inside the jacket, which contains dried air.

The light beam generally interacts with the gas and vapor mixture in the chamber; in some instances this increases the droplet background, worsens track quality, and reduces the ion-load limit. When a chamber is illuminated by means of a gas-discharge lamp, neutral condensation nuclei may appear.[14] Because of their long lifetimes the neutral nuclei can accumulate in the chamber; the droplet background then gradually grows and in the case of prolonged operation of the chamber can become so large that the recording of ionizing particles is impeded. Unfiltered pulsed illumination of a chamber can sometimes prevent normal operation for a long period or lead to total loss of sensitivity.

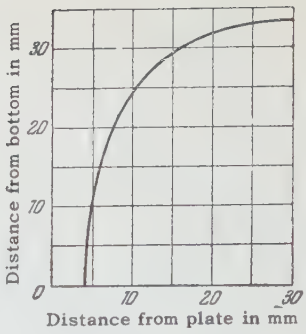


FIG. 15. Profile of sensitive layer close to a plate inside a diffusion chamber.

Light also heats the chamber walls, thus changing the temperature distribution. When the chamber volume is illuminated uniformly and heating of the walls is prevented, prolonged continuous illumination by means of an incandescent lamp does not appreciably affect normal operation. For example, a chamber used to project ionizing particle tracks on a screen functions continuously with a light flux of 2000 lumens.[30]

e) Particle-absorbing Plates in Diffusion Chambers

For the purpose of investigating interactions between high-energy particles and matter it is sometimes desirable to place absorbing plates inside a diffusion chamber. The presence of a solid body inside the chamber and the resulting changed conditions for the transfer of heat and matter can seriously affect the operation of the chamber. A thin plate which is a poor heat conductor has little effect on the temperature distribution in the chamber volume. The temperature distribution on the plate will be that of the chamber gas. Similar conditions exist when composite plates are used whose layers are thermally insulated from each other. In reference 36 it is shown that a lead absorber constructed in this manner does not disturb the thermal regime of a chamber.

When simple plates are used, made of materials that are good heat conductors, the temperature distribution on the plates is necessarily determined by the location of the sources of heat and cold. The plates then have a significant effect on the volume temperature distribution.

Supersaturation cannot exceed unity on the surface of a solid body; an insensitive zone is thus produced close to the plate. This is the principal obstacle which limits the use of plates inside diffusion chambers.

In reference 37 different ways of positioning plates were studied as well as the factors that affect the size of the insensitive zone close to a plate. The size of the insensitive zone was determined by means of alpha and beta-particle sources. An analysis of a track photograph ob-

tained in this work shows that the insensitive zone has the shape represented in Fig. 15. The profile of the sensitive region can be obtained by photographing the chamber volume immediately following expansion or a gas discharge.[8] When the chamber is illuminated by a narrow light beam the photograph clearly shows an insensitive zone close to the walls because of the absence of fog.

The size of the insensitive zone is determined by the vapor flux. With increase of the latter the height of the sensitive zone increases and the insensitive zone is reduced. The character of gas movement within the chamber has a significant influence on the size of the insensitive zone, which can be greatly reduced (to 1 or 2 mm) near the surface of a solid through the flow of gas containing a large amount of vapor at a higher temperature than the surface. This is easily accomplished in a diffusion chamber with a large horizontal component of the temperature gradient.

8. CONCLUSION

Diffusion chambers possess a number of advantages over conventional cloud chambers and are thus replacing the latter at present for some purposes such as work with particle accelerators. Diffusion chambers permit great extension of this method of recording particles. For example, a diffusion chamber can record all ionization events occurring in its sensitive zone and can thus be used for absolute measurements of alpha and beta activities of the order of 10^{-12} Curie.[38] The continuous operation of diffusion chambers makes it easy to record very rare events that are accompanied by the production of strongly ionizing particles.

Diffusion chambers can be used successfully to study gas discharges, the behavior of charged droplets in electric and gravitational fields, as well as the formation and growth of droplets and of ice crystals on different kinds of condensation nuclei. They can also be used to investigate light scattering on droplets[39] and the behavior of a fog or mist-filled region in gravitational and electric fields.[40]

The simple construction and reliable operation of diffusion chambers has also lead to their widespread use for instructional purposes as a demonstration instrument that permits continuous observation of ionizing particle tracks.[30, 41] Diffusion chambers are being used extensively at present to investigate interactions between accelerated particles and matter.[14, 26] They will continue to be important despite the development of such promising instruments as scintillation detectors and bubble chambers.

[1] R. E. Vollrath, Rev. Sci. Instr. **7**, 409 (1936).

[2] A. Langsdorf, Phys. Rev. **49**, 422 (1936).

[3] A. G. Amelin, Теоретические основы образования тумана в химических производствах (Theoretical Foundations of Fog Production in Chemical Manufacturing). 1951.

[4] A. Langsdorf, Rev. Sci. Instr. **10**, 91 (1939).

[5] C. Succi and G. Tagliaferri, Nuovo cimento **9**, 1092 (1952).

[6] Argan, Angelo, and Gigli, Nuovo cimento **1**, 761 (1955).

[7] V. K. Lyapidevskii and Yu. A. Shcherbakov, J. Exptl. Theoret. Phys. (U.S.S.R.) **27**, 103 (1954).

[8] V. K. Lyapidevskii, Dissertation, Moscow Physics and Engineering Institute, 1956.

[9] V. K. Lyapidevskii and O. P. Astakhov, Приборы и техника эксперимента (Instruments and Measurement Engineering) No. 1, 43 (1957).

[10] H. L. Morrison and G. J. Plain, Rev. Sci. Instr. **23**, 607 (1952).

[11] V. K. Lyapidevskii and V. M. Martishin, Приборы и техника эксперимента (Instruments and Measurement Engineering) No. 1, 48 (1957).

[12] V. K. Lyapidevskii J. Exptl. Theoret. Phys. (U.S.S.R.) **29**, 263 (1955), Soviet Phys. JETP **2**, 346 (1956).

[13] V. K. Lyapidevskii, Report to the All-Union Scientific and Engineering Conference on the Use of Radioactive and Stable Isotopes and of Radiation in the National Economy and Science, 1957.

[14] R. M. Sulyaev, Dissertation, Moscow Physics and Engineering Institute, 1956.

[15] O. H. Weddle and C. E. Nielsen, Phys. Rev. **81**, 324 (1951).

[16] Nielsen, Needels and Weddle, Rev. Sci. Instr. **22**, 673 (1951).

[17] P. Harteck and G. Hertz, Naturwissenschaften **39**, 206 (1952).

[18] V. K. Lyapidevskii and O. P. Astakhov, Приборы и техника эксперимента (Instruments and Measurement Engineering) No. 3, 39 (1956).

[19] Miller, Fowler, and Shutt, Rev. Sci. Instr. **22**, 280 (1951).

[20] W. J. Choyke and C. E. Nielsen, Rev. Sci. Instr. **23**, 6, 307 (1952).

[21] R. P. Shutt, Rev. Sci. Instr. **22**, 730 (1951).

[22] A. R. Bevan, J. Sci. Instr. **31**, 45 (1954).

[23] Komar, Stabnikov, and Yashin, Dokl. Akad. Nauk SSSR **108**, 1, 66 (1956), Soviet Phys. "Doklady" **1**, 260 (1956).

[24] V. K. Lyapidevskii, Приборы и техника эксперимента (in press).

[25] Block, Broun, and Slaughter, Phys. Rev. **86**, 583 (1952).

[26] E. C. Fowler, W. B. Fowler, Shutt, Thorndike and Whittemore, Phys. Rev. **91**, 135 (1953).

[27] D. C. Marshall and D. Broun, Rev. Sci. Instr. **24**, 9, 881 (1953).

[28] Alston, Crewe and Evans, Rev. Sci. Instr. **25**, 6, 547 (1954).

[29] Miller, Fowler and Shutt, Rev. Sci. Instr. **22**, 280 (1951).

[30] V. K. Lyapidevskii, **Приборы и техника эксперимента** (Instruments and Measurement Engineering) No. 3, 103 (1957).

[31] Kozodaev, Sulyaev, Filippov, and Shcherbakov, Dokl. Akad. Nauk SSSR **107**, 236 (1956), Soviet Phys. "Doklady" **1**, 171 (1956).

[32] V. K. Lyapidevskii, Author's Cert. (Patent) No. 100, 754 (1952).

[33] V. K. Lyapidevskii, **Приборы и техника эксперимента** (Instruments and Measurement Engineering) No. 3, 40 (1956).

[34] E. M. Munyon and J. D. Kurbatov, Rev. Sci. Instr. **23**, 380 (1952).

[35] W. J. Choyke and C. E. Nielsen, Rev. Sci. Instr. **23**, 307 (1952).

[36] A. Rogozinski, J. phys. et radium **16**, No. 2, 166 (1955).

[37] O. Gordeev, Thesis, Moscow Physics and Engineering Institute, 1955.

[38] V. K. Lyapidevskii, **Приборы и техника эксперимента** (Instruments and Measurement Engineering) No. 2, 37 (1957).

[39] R. L. Lander and C. E. Nielsen, Rev. Sci. Instr. **24**, 1, 20 (1953).

[40] V. K. Lyapidevskii, J. Exptl. Theoret. Phys. (U.S.S.R.) **30**, 399 (1956), Soviet Phys. JETP **3**, 287 (1956).

[41] V. K. Lyapidevskii, **Физика в школе** (School Physics) No. 1, 47 (1953).

Translated by I. Emin

New Instruments and Measurement Methods

HIGH-INTENSITY PULSED LIGHT SOURCES

M. P. VANYUKOV and A. A. MAK

Usp. Fiz. Nauk **66**, 301-329 (October, 1958)

INTRODUCTION

IN many problems involving high-speed photography, stroboscopic observation of periodic processes, optical location, and investigations of the kinetics of photochemical reactions, photoconductivity, luminescence, etc., it becomes necessary to produce, for a brief time interval, a high level of illumination on the observed, investigated, or photographed object. If optical systems are used to concentrate light beams, high-intensity light sources (in which the light is produced by a body of great brightness) become particularly important.

The brightness of a source of illumination is determined primarily by the temperature of the glowing substance. The maximum temperatures that can be obtained in solids are ~ 4000°K, those in liquids are ~ 6000°K. Only gases or vapors can exist at higher temperatures. As the temperatures increase, the degree of ionization of a gas increases and the gas turns into plasma. The plasma temperature may reach very high values, reaching 10^7 to 10^9 °K in thermonuclear reactions.

It must be borne in mind that when light sources with glowing plasma are considered, they are volume sources and that the power radiated per unit surface of the glowing body into a unit solid angle per unit spectral interval, called the spectral energy density b_λ, is determined by the following expression

$$b_\lambda = b_\lambda^* (1 - e^{-k_\lambda l}),$$

where b_λ^* is the black body spectral energy density at a temperature T, equal to the temperature of the light source; k_λ is the absorption coefficient of the plasma for radiation of wavelength λ and l is the thickness of the glowing body (the plasma temperature is assumed constant over the entire glowing volume).

The foregoing expression shows that the brightness of a plasma light source depends not only on the temperature but also on the thickness of the radiating layer and on the absorption coefficient of the plasma. The maximum brightness that can be obtained at a given temperature is determined by the black-body radiation laws. An example of a high temperature light source that has a relatively low intensity of continuous radiation is the low-pressure, high-power pulsed electric gas discharge, used to effect controllable thermonuclear reactions. A compression of the discharge resulting from the pinch effect produces in the plasma a temperature of several millions of degrees. However, owing to the high transparency of the plasma[1] (the mean free path of the photons in the discharge is usually many times greater than the dimensions of the discharge chamber) this discharge channel is of low brightness.

In this survey we examine methods of producing high-brightness pulsed light sources, and also methods for measuring their emission characteristics.

I. SPARK DISCHARGE IN GASES

One of the most widely used pulsed light sources, in which the glowing body has a high instantaneous brightness, is produced by a spark discharge in high-pressure gases.

The spark-discharge radiation is caused by the retardation of the electrons in the field of the positive ions (free-free transitions), electron-ion recombination (free-bound transitions), and radiation of greatly broadened lines (bound-bound transitions).[2] For hydrogen and hydrogen-like atoms, Unsold[3] obtained an expression for the spectral distribution of the plasma radiation. However, radiation from the powerful spark discharge that is usually used as a source of light is closer in its spectral characteristics to black-body radiation.[4]

As early as in 1852, Talbert[5] has used the flash produced when a capacitor charged to high voltage is discharged through an air gap to photograph fast objects. Recently pulsed spark lamps appeared, in which the discharge is produced in an atmosphere of heavy inert gases (argon, crypton, or xenon) at gas pressures of several atmospheres. Such lamps have a greater light yield than an ordinary spark discharge in air, but also a considerable longer glow time.

During the last two decades low-pressure

pulsed lamps have become widely popular among photographers. These usually are filled with xenon and produce approximately 30 to 40 lumens per watt. However, we shall not consider these lamps, the brightness of which is relatively small.

In this section we shall consider the conditions under which maximum brightness is obtained with a pulsed electric discharge in gases at high pressure, in the case when the glowing body is a current-carrying channel that makes no contact with the walls of the chamber.

I. 1. Methods of Producing High-Intensity Spark Discharges

To produce high brightness in the spark-discharge channel it is necessary above all to insure the maximum possible speed of power intake by the channel. This is obtained by discharging at a sufficiently high voltage, low inductance in the discharge circuit, and high gas pressure.

An idea of the influence of individual parameters of the discharge circuit on the speed of power intake by the discharge gap can be gained from considering the current variation in a damped oscillating discharge. At the beginning of the discharge (during the first half cycle of current), the current i is given with sufficient accuracy by the following expression

$$i \cong U_0 \sqrt{\frac{C}{L}} \sin \omega t,$$

where U_0 is the initial capacitor voltage, C the capacitance and L the inductance of the discharge circuit, $\omega \cong 1/\sqrt{LC}$ the natural frequency of the discharge circuit, and t the time.

The rate of current build up in the discharge, di/dt, determines in first approximation the rate of power intake by the spark gap. Differentiation of the expression for the current shows that at the initial instant of time di/dt depends only on the voltage and inductance of the discharge circuit

$$\left(\frac{di}{dt}\right)_{t=0} = \frac{U_0}{L}.$$

The maximum value of the current in a low-damping circuit is given by

$$i_{max} = U_0 \sqrt{\frac{C}{L}}.$$

The rate of broadening of the channel also affects the instantaneous energy density in the channel. Using discharges in air at pressures of 200 mm Hg and 3 atmos as examples, Gegechkori[6] has shown that the radius of the channel increases more slowly at increased pressure. It was also established[7] that, other conditions being equal, the speed of broadening of the channel varies in different inert gases and diminishes with increasing atomic weight of the gas. The channel diameter d, measured 5 to 8 μsec after the start of the discharge, when its subsequent change becomes insignificant, depends on the capacitor voltage U_0, the molecular weight of the gas M, and its pressure p in the following manner:

$$d = B \sqrt{\frac{U_0}{pM}},$$

where B = 28 for a circuit with C = 1.09 μf and L = 0.6 microhenry (the voltage U_0 is in kilovolts and p is in atmospheres).

Based on the above considerations, the discharge channel, other conditions being equal, is expected to have a higher instantaneous brightness in heavy inert gases than in the lighter ones. Actually, the reverse is true, owing apparently to the fact that the plasma temperature is dependent to a great degree on the specific heat, the heat conduction, and the electric conductivity of the gas, which at high temperatures depend in turn on the ionization potential rather than on the difference in the diameters of the current-conducting channels, a difference caused by the unequal rate of their broadening.

If pulsed lamps are made in sealed glass or quartz bulbs, the pressure seldom exceeds 5 to 8 atmos, in order to reduce the danger of the bulb bursting. To produce a discharge at higher pressures it is possible to use dismountable pulsed lamps.[4,8,9,11,14] Vanyukov et al.[8] describe the construction of a lamp that operates at 10 atmos. Data are also published on light from a spark discharge in hydrogen at pressures up to 95 atmos.[9]

To construct high intensity spark light sources it becomes necessary to develop special low-inductance, high-voltage capacitors and to devise the most suitable methods for connecting these capacitors to the discharge gap. In such light sources the capacitor and the pulsed lamp proper usually comprise an integral unit. To reduce the inductance of the discharge circuit it is necessary to aim for such an arrangement of the circuit elements, at which the magnetic field formed by the current is contained in the smallest possible volume.

Data on low-inductance capacitors and discharge circuits are reported in references 8 to 16. By way of an example, we cite certain characteristics of low-inductance capacitors used in high-speed photography. Fayole and Naslin[13] report on a circuit with C = 0.004 μf and an L = 0.006

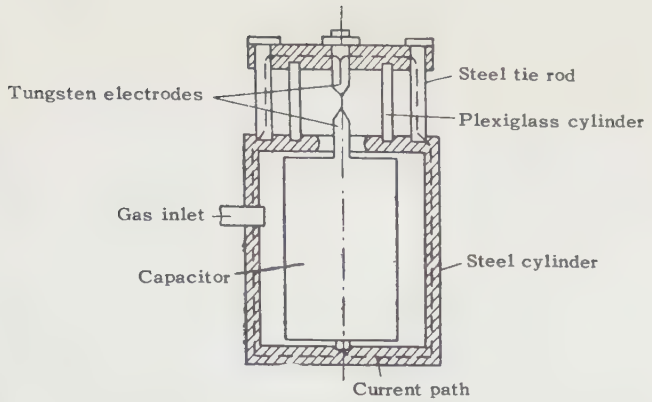


FIG. 1. Diagram of a dismountable pulsed lamp with capacitor placed inside the lamp housing (C = 0.1 μf, U_0 = 12 kv, L = 0.03 microhenry).

microhenry, obtained with specially designed ceramic capacitors. A feature of another type of low-inductance 0.01-μf capacitor is that it produces a 2-joule aperiodic discharge which is completed within practically 10^{-7} seconds.[13]

Let us examine now several methods used to obtain a low-inductance connection between the capacitor and the spark-discharge gap. Figure 1 shows schematically the arrangement of a dismountable spark lamp, proposed by Früngel.[11] A bushing-type capacitor is placed inside the metal case of the lamp, whose walls serve to carry the current. Light from the spark is transmitted through a plexiglass cylinder. Kovaszney[12] describes a circuit in which a block of 6 capacitors

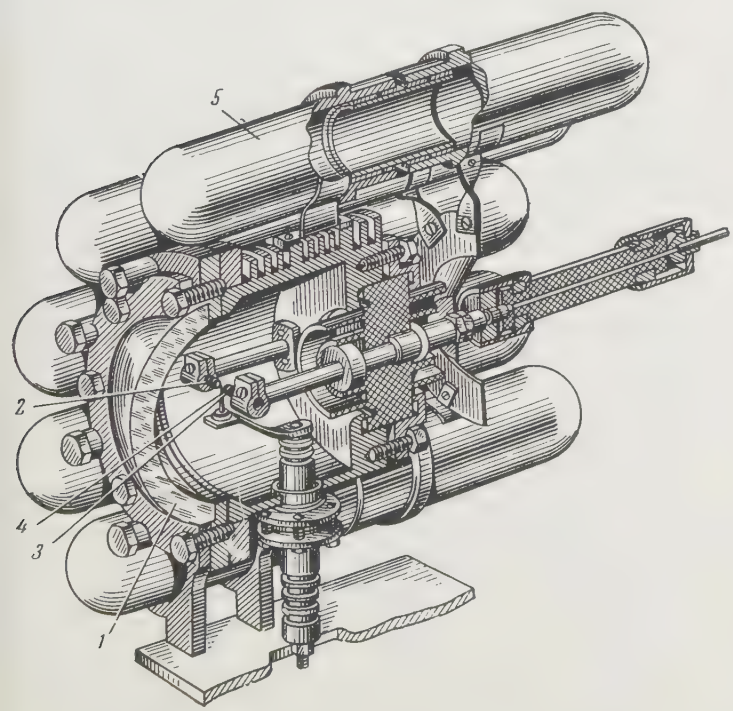


FIG. 3. Dismountable pulse lamp; 1) exit window, 2 and 3) principal electrodes, 4) control electrode 5) ceramic capacitors.

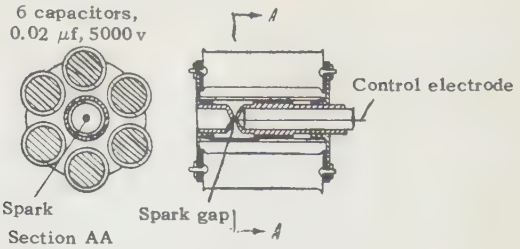


FIG. 2. Diagram of a discharge circuit with capacitors in parallel.

connected in parallel surrounds the discharge gap, as shown in Fig. 2. The capacitors are interconnected with metallic plates. Light passes through a 0.5 mm diameter hole in the electrode. A shortcoming of such a light source is the small angle of light emergence. Vanyukov et al.[8] developed a dismountable pulsed lamp with ceramic capacitors, which emits light at an angle of 120°. In this lamp (Fig. 3) a capacitor bank with a total capacitance of 0.022 μf is charged to 28 kv and then discharged into a circuit with an inductance of 0.06 microhenry. The lamp can withstand prolonged operation at an average power up to 4 kw.

High-intensity short flashes can be obtained by discharging a long line with distributed capacitance and inductance, charged to a high voltage, through a spark gap.[17-19]

We know that if such a line is discharged into a matched load, it delivers to the load a rectangular current pulse, whose duration is equal to twice the time that it takes the wave to travel through the line. In view of the fact that the spark gap does not have a constant resistance during the discharge, it cannot be matched to the characteristic impedance of the line. However, it is advisable in any case to use a line with a low characteristic impedance, for it is well known that after the breakdown the resistance of the discharge gap is a fraction of an ohm. Fitzpatrick et al.[19] used as such a line a segment of a coaxial cable with barium titanate as the dielectric. In view of the high dielectric constant of the barium titanate ($\epsilon \cong 1000$) the characteristic impedance of the cable is only 1.2 to 1.5 ohms. The cable is made in the form of a section of a hollow cylinder with inside and outside diameters 2.5 and 5 cm. The inside and outside surfaces are silver

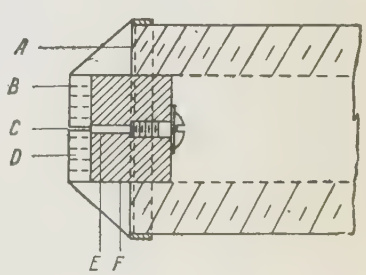


FIG. 4. Discharge circuit with coaxial cable; A) $BaTiO_3$ cylinder, B) steel electrode, E) tungsten electrode, C) discharge gap 0.3 mm in diameter, D and F) insulators.

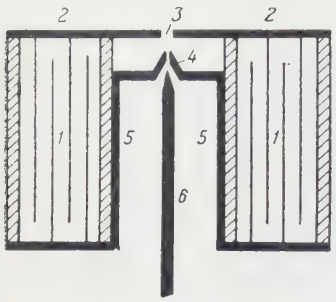


FIG. 5. Discharge circuit with coaxial capacitor, 1) coaxial capacitor, 2) metallic plate with aperture, 3, 4) electrodes, 5) metal tube, 6) control electrode.

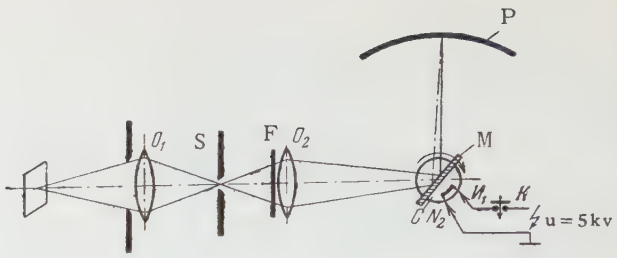


FIG. 6. Optical diagram of photochronograph used to record the glow of explosions.

plated (Fig. 4) and are connected to the discharge gap. A cable 16.5 cm. long is used to obtain light flashes of 10^{-7} sec duration.

The light source described above does not produce, however, high-intensity flashes for to increase the capacitance of the discharge circuit it becomes necessary to increase the length of the cable, which in turn increases the duration of the discharge. This difficulty is overcome by using as the working capacitance a short but multilayered coaxial cable, in which the corresponding electrodes are interconnected in paralle. A schematic diagram of a discharge circuit with such a capacitor, as proposed by Fischer,[15] is shown in Fig. 5. With a capacitance of 0.1 μf and 3 kv applied, the inductance of this circuit is merely 0.004 microhenry. Fischer also reports[16] the development of a toroidal capacitor for a higher voltage ($C = 0.6$ μf, $U_0 = 60$ kv, $L = 0.14$ microhenry), so constructed that it surrounds the discharge gap coaxially.

A low-inductance discharge gap for the production of large pulsed currents has been developed by V. S. Komel'kov and G. N. Aretov.[20] They used flat massive connections for the capacitors and reduced to a minimum the distance between bus bars; they thus succeeded in constructing a discharge circuit comprising 48 capacitors, with a total capacitance of approximately 134 μf at a working voltage of 50 kv, and with a total inductance (disregarding the load inductance) of nearly 0.025 microhenry. The maximum current in such a circuit reaches 2.1×10^6 amp with $(di/dt)_{max} = 2 \times 10^{12}$ amp/sec.

We have already mentioned the principal methods of obtaining spark discharges in which a high rate of power intake is produced by reducing the inductance of the discharge circuit to the limit. Another way of accomplishing the same purpose is to produce the discharge at the highest possible voltage. This method was used by Magan and Woerner[21] to produce a spark discharge with a capacitor of 300 $\mu\mu$f charged to 450 kv. The discharge capacitance was that of the electrodes themselves. Owing to the great length of discharge gap a high voltage discharge has a good light efficiency. The radiation from such a discharge was studied in a special chamber, filled with argon and nitrogen at pressures up to 20 atmos. Depending on the pressure, the duration of the light flashes varied from 0.3 to 1.0 μ sec.

I. 2. Methods of Measuring the Brightness and Temperature of Pulsed Light Sources

The measurement of the instantaneous brightness characteristics of pulsed light sources involves considerable experimental difficulties. These are due primarily to the short duration of the radiation process. Fischer[15] and Vanyukov et al.[22] investigated light flashes with a front duration of approximately 10^{-7} seconds. Consequently the apparatus intended for recording the time variation of the radiation intensity of short light flashes should have a very small time constant. The apparatus used in some cases has a time resolution of 10^{-7} to 5×10^{-8} seconds.[22,23,24]

The apparatus used to measure the absolute values of the brightness of pulsed sources is usually calibrated with the aid of standard sources of known brightness or color temperature. The sources employed are incandescent lamps,[22,23,25,26] zirconium lamps,[15] and the sun.[24,27] Of all these sources, the sun has the highest intensity, but its temperature is approximately 6000° K, while that of pulsed sources frequently exceeds 20,000 or 30,000° K. This reduces the measurement accuracy, which depends on the ratio of the temperatures of the standard and investigated sources.[24]

The first measurements of the temperature of pulsed sources of light were made by Anderson and Smith.[25,28] The measurement apparatus, described in reference 25, consists of a monochromator, thermocouple, and galvanometer. The apparatus is calibrated with a comparison lamp. The galvanometer operates as a ballistic instrument when recording the pulsed signal. The brightness of the source, averaged over the flash duration, is determined from the reading of the galvanometer and from the effective duration of the light flash. The

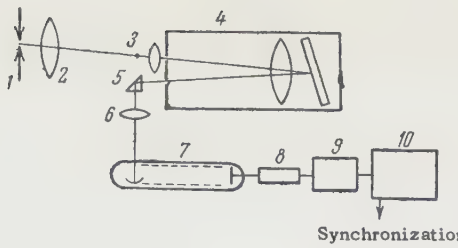


FIG. 7. Schematic diagram of the photoelectric setup used to measure the brightness of a spark-discharge channel.

duration of the light flash is determined by the author from photographs of a mirror scan of the flash. Considering that the effective duration of the signal was estimated rather roughly, the accuracy of brightness measurement by this method is probably not very high.

Later on the brightness and the temperature were measured either by photographic[24,27] or by photoelectric methods.[4,11,15,22,23,26,29,30,31,32]

As an example of the use the photographic method for measurement of instantaneous values of the temperature, we refer to a paper by I. Sh. Model'[24] on the determination of the temperature reached in shock compression of gases.

The optical system of the apparatus, comprising a high speed photochronograph with a rotating mirror, is shown in Fig. 6. Slit S passes over part of the image of the investigated event which is focused in the plane of the slit by objective O_1. The image of the slit is projected by objective O_2 on the photographic plate P after being reflected by the rotating mirror M. Filter F comprises jointly with the photographic film a simple monochromatizing system ($\Delta\lambda \approx 900$A). The exposure time can be adjusted from 10^{-5} to 10^{-7} seconds. The same setup was used to photograph the blackening markers produced by the standard, in this case the sum. To permit the use of the reciprocity law, the film exposure time did not exceed 10^{-5} seconds when calibrating with the standard. This was accomplished by using a shutter of the focal plane type. The temperature was measured in this setup with an accuracy of 6 to 20 percent, depending on the absolute values of the measured temperature.

It must be noted that the photographic procedure for brightness measurement has several shortcomings, the major one being the low sensitivity. Furthermore, the processing of the results is complicated and labor consuming.

The photoelectric method of brightness measurement is much more sensitive. The radiation receivers usually used in this method have a very high time resolution (up to 10^{-8} or 10^{-9} sec), and the processing of the measurement results is quite simple and takes little time. In the first experiments the time variation of the light intensity of the flashes was recorded photoelectrically. The brightness of the channel was determined from these data with allowance for the effective area of the glowing substance.[11,30] A shortcoming of the method is that it can be used to determine only the brightness averaged over the channel. Furthermore, its accuracy is low owing to errors in the determination of the area of the glowing substance.

The first to measure directly the brightness of a discharge channel in pulsed lamps were Vanyukov et al.,[22] who used the photoelectric method. The schematic diagram of the measuring setup is shown in Fig. 7. A small portion of the glowing substance of the lamp 1 is projected by lens 2 on the plane of the entrance slit of monochromator 4. The bandwidth of the monochromator is 26 A. The uniformly illuminated objective of the monochromator is projected on the cathode of the photomultiplier 7 by lens 6. The output signal of the photomultiplier passes through delay line 8 and is applied to amplifier 9, and then to the plates of oscillograph 10. The time resolution of the photoelectric system amounted to 7×10^{-8} seconds. The sensitivity of the system was regulated by introducing neutral attenuation filters. Owing to the spatial instability of the discharge channel, it was impossible to record with this method the time variation of the brightness of the brightest portion of the glowing substance within a single discharge. Consequently 100 to 150 oscillograms of the flashes were photographed and the oscillograms with maximum signal amplitudes were selected.

The sensitivity of the apparatus was calibrated with the aid of an incandescent lamp ($T = 2073°K$). A rotating-mirror arrangement made it possible to use modulated light for the calibration. The procedure described produces wavelength vs. spectral energy density curves for the brightest portion of the glowing substance for various instants of time. These data afford a most complete description of the brightness and can be used to determine the visual brightness and the effective brightnesses for other selective radiation receivers (photographic film, photocells, etc.).

An analogous procedure for the measurement of the brightness of pulsed sources of light is used in references 23, 26, 29, 31 and 32.

In another interesting procedure, recently described by Fischer,[15] the instantaneous value of the visual brightness of pulsed sources is determined directly. The radiation is picked up by a photomultiplier, the output signal from which is recorded with an oscillograph. Narrow spectral filters are used, maximum transmission at

$\lambda = 5450$ A. The standard used is a dc zirconium lamp with a visual brightness of 4050 c/cm^2. A rotating disk transforms the dc-light signal from the standard source into a pulsed signal, whose duration is comparable in length with that of the pulsed source. Neutral attenuation filters are used to reduce the amplitude of the signals from the pulsed source. The visual brightness B of the investigated light source is obtained from the expression

$$B = B^* \cdot \frac{S}{S^*} \cdot \frac{f}{D} \text{ c/cm}^2,$$

where B* is the brightness of the standard lamp, S the amplitude of the pulsed-lamp signal, S* is the amplitude of the standard-lamp signal, D is the transparency of the attenuating filters, and f is a color factor that takes into account the difference in the spectral emission of the investigated and standard lamps.

The color factor f is determined by the author from the following relations:

$$f = \frac{U^*}{U} \cdot \frac{b}{b^*};$$

$$U = \int E_\lambda T_\lambda Z_\lambda \, d\lambda;$$

$$U^* = \int E_\lambda^* T_\lambda Z_\lambda \, d\lambda;$$

$$b = \int E_\lambda V_\lambda \, d\lambda;$$

$$b^* = \int E_\lambda^* V_\lambda \, d\lambda,$$

where E_λ is the relative spectral radiation density of the pulsed lamp, E_λ^* the relative spectral radiation density of the standard lamp, T_λ the transmission of the filters, Z_λ the relative spectral sensitivity of the photomultiplier, and V_λ the relative spectral eye sensitivity.

In the author's opinion, the error in f did not exceed 5 percent, and the error in the visual brightness B did not exceed 10 percent (concerning the results obtained by the author, see Sec. I.3).

To understand the physical processes in the spark-discharge channel, and in particular to ascertain whether it is possible to increase its brightness, it is important to determine the temperature of the discharge channel. Since the processes connected with the spark discharge are highly non-stationary, such terms as temperature, equilibrium state, etc. are evidently not applicable to a light-emitting plasma.[33] References 33 to 37 are devoted to the problems that are raised thereby.

S. L. Mandel'shtam and N. K. Sukhodrev have shown[36] that the excited states of the atoms in the spark-discharge channel have a Boltzmann distribution, while the ionization is given by the Sach formula. In both cases the role of temperature is played by the electron temperature. The excitation distribution becomes stationary within approximately 10^{-10} sec, and the ionization becomes stationary within 10^{-7} sec. The time required for the electron and gas temperatures to become equalized is also approximately 10^{-7} sec. It is shown in reference 37 that in the plasma of the spark-discharge channel the relation between the radiating and absorbing ability of the radiator is given by Kirchhoff's law, provided the temperature of the latter is taken to be the electron temperature. It was also shown[38,39] that the Boltzmann distribution holds for the excited atoms in the plasma of the spark-discharge channel if the gas pressure is not less than 100 mm Hg. It was established in references 22, 23, 29, 31, and 32 that at certain discharge conditions the spectral distribution of the emission of the spark-discharge channel follows roughly that given by Planck for a black body.

It has thus been established at present that the plasma of the spark-discharge channel can be assigned a definite temperature if the time exceeds 10^{-7} sec. For time durations less than 10^{-7} seconds the emission of the spark discharge is determined by the electron temperature of the plasma. This must be borne in mind in cases when the plasma temperature is determined from the emission of the spark-discharge channel.

Having made these remarks we can proceed to a brief description of the existing methods for determining the temperature of the glowing substance in pulsed-light sources. The brightness method of determining the temperature was used in references 23, 24, 29 and 32. In this method the brightness of the investigated body is compared with that of a standard of known temperature, determined from the following expression[24]

$$\frac{(e^{\frac{c_2}{\lambda T}} - 1)}{(e^{\frac{c_2}{\lambda T_0}} - 1)} = \tau a,$$

where T_0 is the temperature of the standard, T the temperature of the investigated light source, λ the effective wavelength of the employed spectral interval, τ the ratio of the brightness of the standard to the brightness of the investigated body, a the absorption coefficient of the investigated body, and $c_2 = 1.438$ cm-deg.

This method is used most frequently in those cases when it is established somehow that the absorption coefficient of the glowing body is unity.[23,29,32] Under certain conditions (brightness "saturation") the temperature can be measured

by the brightness method using not only continuous but also line-spectrum radiation.[23]

Babushkin[40] and Glaser[41] use a color method for determining the temperature, in which the spectral distribution of the unknown body and of a black body are compared. It must be kept in mind, however, that this method is less accurate than the brightness method at high temperatures (> 10,000° K).[24]

If the plasma of the discharge channel is transparent the temperature can be determined from the ratio of the intensities of two spectral lines.[37,39,42,43] The temperatures are determined from the relation

$$\frac{I_1}{I_2} = \frac{A_1 g_1 \nu_1}{A_2 g_2 \nu_2} e^{\frac{E_2 - E}{kT}},$$

where I_1 and I_2 are the line intensities, A_1 and A_2 are the transition probabilities, g_1 and g_2 are the statistical weighting factors of the above terms, ν_1 and ν_2 are the frequencies of the lines, E_1 and E_2 are the excitation energies of the above levels, k the Boltzmann constant, and T the unknown temperature.

It must be borne in mind that to determine the maximum temperature in the discharge channel it is necessary to use lines of highly-ionized atoms, and also to time-scan the spectrum.[42,43]

Various non-optical methods of temperature determinations are also used. Thus, in references 28 and 40 the temperature is determined from the sound velocity, using the following relation

$$v = \sqrt{\frac{\gamma RT}{m}},$$

where v is the sound velocity in a medium of temperature T, R is the gas constant, γ is the ratio of the specific heats, and m is the molecular weight.

Since the molecular weight is also a variable quantity at high temperatures, this equation must be solved simultaneously with the Sach equation.

V. S. Komel'kov and D. S. Parfenov[44] use the following expression from the pinch-effect theory[45] to determine the plasma temperature of the spark-discharge channel

$$2NkT = I^2,$$

where I is the current of the discharge, N the total number of particles per centimeter length of the channel, k Boltzmann's constant, and T the average gas temperature in the channel.

Simultaneous solution of this equation with the Sach equation yield the temperature T.

Several other methods for measurement of the spark-discharge channel temperature are discussed in reference 41.

I. 3. Maximum Brightness Obtained with the Aid of a Spark Discharge in Gases

The first systematic investigations of optical characteristics of the spark discharge in an atmosphere of various gases were made by Glaser in 1950-1951.[4,41] In particular, he investigated the brightness of the spark-discharge channel in an argon atmosphere by mirror scanning. By photometry of the resultant photographs Glaser has established that the maximum instantaneous brightness along the axis of the discharge channel increases with increasing power supply to the channel only up to a certain limit.[4,41] Further increase in the power leaves the brightness unchanged at a value corresponding to the brightness of a black body at a temperature of approximately 40,000° K.[41]

K. S. Vul'fson, I. Sh. Libin, and F. A. Charnaya[46] used an analogous procedure for a systematic investigation of the brightness of the spark-discharge channel in atmospheres of argon, krypton, and xenon at various gas pressures and at various discharge energies. The authors have also reached the conclusion that the brightness of the discharge channel in these gases reaches a limit with increasing voltage. It was established that the voltage at which the limiting brightness is reached depends on the type of gas and on its pressure. Thus, for example, at a gas pressure of two atmospheres the limiting brightness in xenon is reached at a discharge-circuit voltage of 5 kv, in krypton at 6 kv, and in argon only at 10 kv. As the pressure increases, the maximum brightness is reached at lower voltages.

The effect of brightness saturation was investigated by Vanyukov, Mak, and Parazinskaya,[22] and also by Vanyukov and Mak,[23] who used a photoelectric procedure to register the time variation of the spectral brightness density of the central portion of the channel in energy units. Using this procedure, Vanyukov et al.[22] investigated the continuous radiation from a spark discharge in argon and xenon at pressures of 4 to 5 atmos in the spectral region from 4000 to 9000 A, with a time resolution of 0.07 μsec and with a spectral resolution of approximately 26 A. They also investigated the influence of the inductance of the discharge circuit on the brightness of the channel. A change in the inductance of the discharge circuit affects greatly the energy intake by the spark-discharge channel,[47] and therefore a study of the affect of the inductance

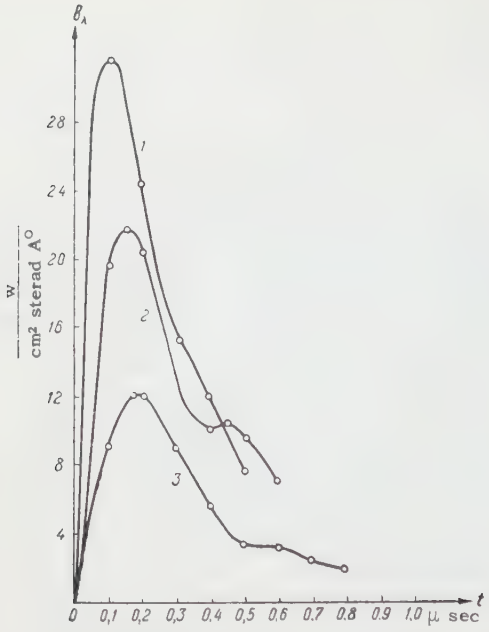


FIG. 8. Time dependence of b_λ of the discharge channel in argon for various values of circuit inductance; 1) 0.12 microhenry, 2) 0.6 microhenry, 3) 1.1 microhenry; λ = 4680 A, U_0 = 12 kv, C = 0.011 μf.

on the spark-channel brightness is of substantial interest.

Figures 8 and 9 show typical curves for the dependence of the spectral brightness density b_λ in the central portion of the channel on the time t for various values of the inductance of the discharge circuit for discharges in argon (Fig. 8) and in xenon (Fig. 9). It was established that the brightness saturation is easiest to obtain in the long-wave portion of the spectrum. As the radiation wavelength decreases, the maximum brightness is obtained at lower inductance values. By way of an example, Fig. 10 shows the variation of the maximum instantaneous spectral brightness density of the spark-discharge channel in an argon atmosphere with the inductance of the discharge circuit for various wavelengths. In the case of discharge in xenon maximum channel brightness was obtained over the entire investigated spectral region even when the maximum circuit inductance was used.

It was also established that when the maximum brightness is reached, the radiation from the discharge gap closely approximates the radiation of a black body with a temperature determined from the value of the spectral density of the discharge-channel brightness. Figure 11 shows the spectral variation of b_λ of a discharge in xenon for various instants of time lapsed since the beginning of the discharge. The figure shows that at the instant t = 0.1 μsec the discharge radiation corresponds to that of a black body at T = 27,000°K.

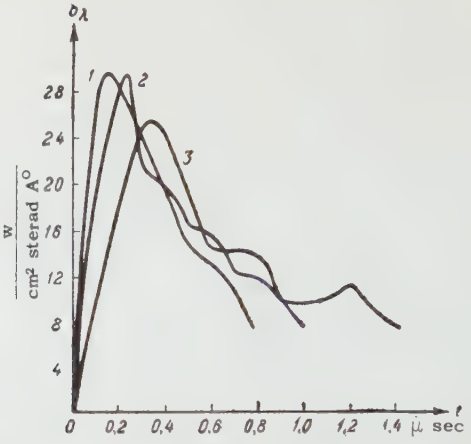


FIG. 9. Time dependence of b_λ of the discharge channel in xenon for various values of circuit inductances: 1) 0.12 microhenry, 2) 0.6 microhenry, 3)1.1 microhenry; λ = 4490 A, U_0 = 12 kv, C = 0.011 μf.

The maximum brightness of the spark-discharge channel in argon exceeds the maximum brightness of the channel in xenon, although the brightness of xenon lamps is higher than those of argon at slow rates of energy intake by the discharge channel (Fig. 12).

Vanyukov and Mak[23] extended the investigation of the continuous background of the brightness spectrum into the ultraviolet region of the spectrum (to 2300 A), and also investigated the line-spectrum radiation from ionized gases in which a discharge was produced. This investigation has shown that with increasing discharge-circuit in-

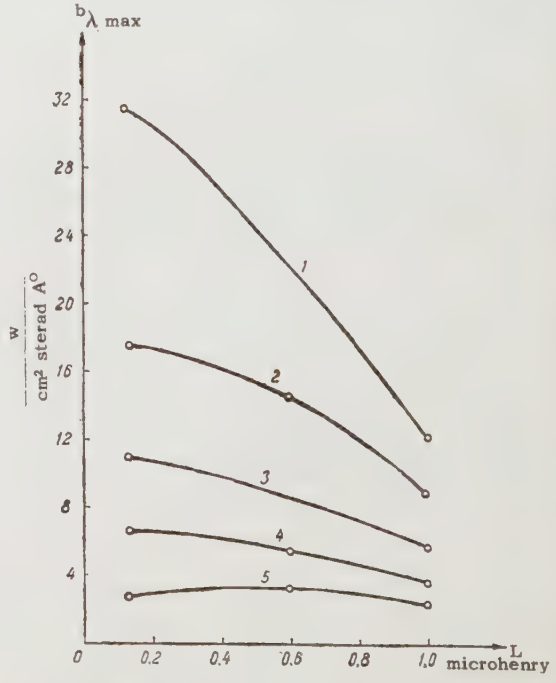


FIG. 10. Dependence of b_λ of the discharge channel in argon on the discharge-circuit inductance for various wavelengths: 1) 4,680 A; 2) 5540 A; 3) 6520 A; 4) 7230 A; 5) 8870 A.

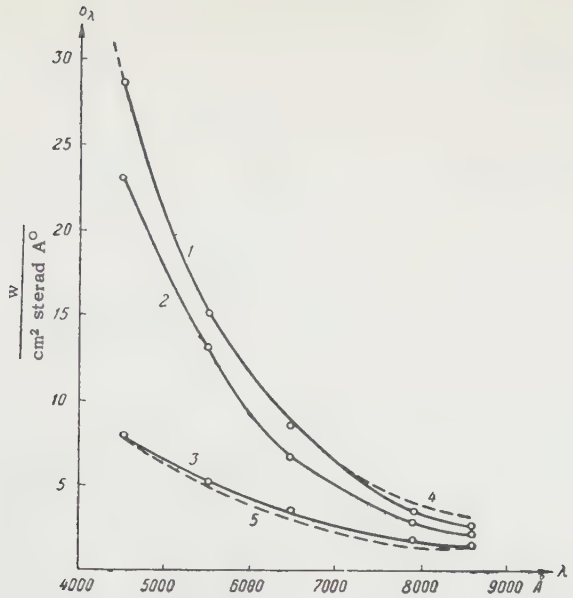


FIG. 11. Dependence of b_λ of the discharge channel in xenon on the wavelength for various instants of time: 1) 0.1 μ sec, 2) 0.3 μ sec, 3) 0.8 μ sec, 4 and 5) b_λ of a black body at 27,500°K and 14,500°K, respectively.

ductance the line-spectrum radiation intensity reaches a maximum more rapidly than the intensity of the continuous background, owing to the large absorption coefficient of the plasma in the lines. At the same time the temperatures, determined from the maximum value of the spectral brightness density of the continuous and line radiation are in good agreement, thus confirming the conclusion made in reference 22 that as brightness saturation is reached the discharge channel radiates like a black body. The degree of correspondence between the radiation of the spark-discharge channel and the radiation of a black body is illustrated in Fig. 13, which shows the wavelength dependence of the temperature for discharges in argon, xenon, and nitrogen. From the temperature values shown in the diagram it is possible to determine with the aid of Table IV (see Sec. VI) the value of the maximum visual brightness. In accordance with reference 22, it amounts to 11×10^6, 15×10^6, and 21×10^6 c/cm^2 for xenon, argon, and nitrogen respectively.

F.A. Charnaya investigated the dependence of the spark-discharge channel brightness in air,[48] xenon, krypton, argon, oxygen, nitrogen, neon, and helium[26] on the breakdown voltage, on the gas pressure, and on the distance between the electrodes (absolute brightness values were measured in reference 26). In all these gases, with the exception of helium, the author has established that the spark discharge channel reaches maximum brightness. When the discharge is formed in air[48] the maximum

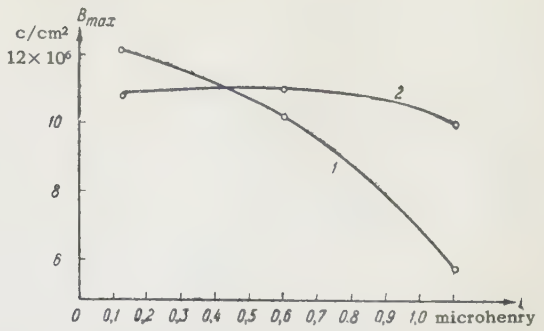


FIG. 12. Dependence of the maximum instantaneous brightness of the channel of the inductance of the discharge circuit: 1) argon, 2) xenon.

brightness was seen to increase with pressure in a range up to 7 atmos. The maximum brightness remained unchanged when the distance between electrodes was varied from 1.5 to 6 mm. It was also found that the maximum brightness of the discharge channel in air exceeds the maximum brightness in argon at the same pressure. Reference 26 also states that as the atomic weight of the gas decreases, the maximum brightness increases thus confirming the results of reference 22 for argon and xenon.

A study of the dependence of the value of the maximum brightness on the pressure of the filling gas[26] has shown that at sufficiently large pressures the maximum brightness is independent of the pressure.

Fischer[15,16,49] has investigated the visual brightness of a spark-discharge channel in air and helium, the discharge being produced in a circuit with very small inductance (L = 0.004 microhenry, C = 0.1 μf). The maximum brightness found in air was 45 to 50 × 10^6 c/cm^2. In helium at p = 35 atmos

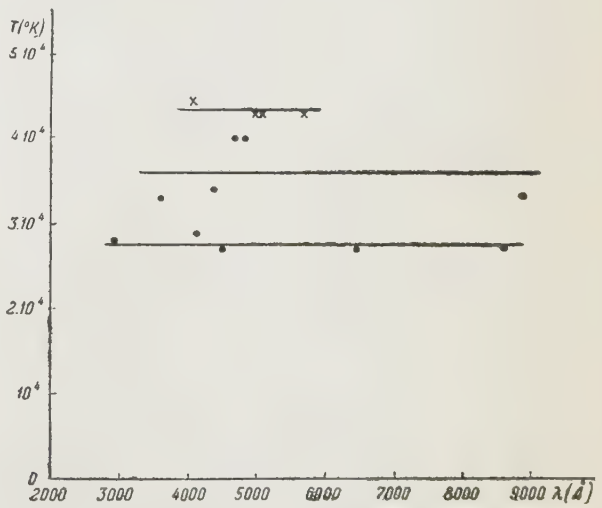


FIG. 13. Temperature of the spark-discharge channel in xenon, (•), argon (○) and nitrogen (×), C = 0.05 μf, L = 0.086 microhenry, U_0 = 12 kv (for argon and xenon), and U_0 = 15 kv (for nitrogen).

TABLE I

Gas	Temperature, °K	Visual brightness, c/cm²	Authors, reference
Xenon		1.7×10^8	Früngel[11]*
	27 000	1.1×10^7	Vanyukov, Mak, and Parazinskaya,[22] Vanyukov and Mak[23]
Argon	140 000	2×10^8	Früngel[11]*
	40 000		Glaser[41]
	35 000	1.5×10^7	Vanyukov and Mak[23]
	37 000		Model'[24]
Air (nitrogen)		5×10^7	Fischer[15]*
	43 000	2.1×10^7	Vanyukov and Mak[23]
	44 500		Charnaya[26]**
		4×10^8	Edgerton and Cathon[50]*
Helium	250 000	1.5×10^8	Fischer[15]*

*In our opinion these results are too high.
**Calculated from the maximum brightness cited in this paper for λ = 9000 A.

and U = 7 kv (maximum discharge current I = 63 kilo amperes, discharge energy W = 68.5 joules) the maximum brightness was on the order of 150×10^6 c/cm², corresponding to a temperature of T = 250,000°K. This is apparently the highest temperature registered in a spark-discharge channel. It must be noted, however, that the maximum brightness for air, measured by the same author, disagrees with the results of reference 23, where it has been established that the temperature of the spark-discharge channel in nitrogen is 43,000°K, corresponding to a maximum visual brightness of 21×10^6 c/cm².* It can therefore be assumed that Fischer's values of brightness and temperature of a spark-discharge channel in helium are too high. There is no doubt, however, of the fact that a spark discharge can produce a much greater brightness in helium than in air.

From an examination measurements that are the most reliable from our point of view, we can conclude that for each gas there exists a limiting brightness, which can be obtained with a spark discharge of sufficient power.

The maximum brightness is reached most easily in gases having low ionization potential.

Table I gives a summary of the values of temperature and of visual brightness obtained with spark discharges in atmospheres of various gases.

*The temperature of the spark-discharge channel in nitrogen, obtained from the value of maximum brightness at λ = 9000 A as found in reference 26, is approximately 44,500°K, which is in good agreement with the results of reference 23.

I. 4. Physical Limitations on the Brightness of the Spark-Discharge Channel

An investigation of the emission of the spark-discharge channel and of the distribution of the brightness of the glowing substance, and measurement of the transmission coefficient of the discharge-channel plasma, have shown that in the brightness saturation mode the discharge channel is opaque and radiates as a black body (in those portions of the spectrum where saturation is reached).[4,15,22,23] In connection with this, a hypothesis has been set forth in references 15 and 26, that the existance of a maximum brightness is related to the opaqueness of the spark-discharge channel at high temperatures.

A similar point of view was set forth by Ya. B. Zel'dovich[51,53] and developed by Yu. P. Raizer[52] in connection with the radiation from strong shock waves. Taking into account the strong dependence of the radiation absorption coefficient on the temperature (the Kramers formula was used to estimate the absorption coefficients of air at high temperatures), the authors have concluded that at high temperatures there is formed in front of the shock wave a heated zone that begins to shield the radiation travelling from the front of the wave. According to calculations,[52] the heated layer of air shields completely the outward radiation even when the temperature behind the shock wave front

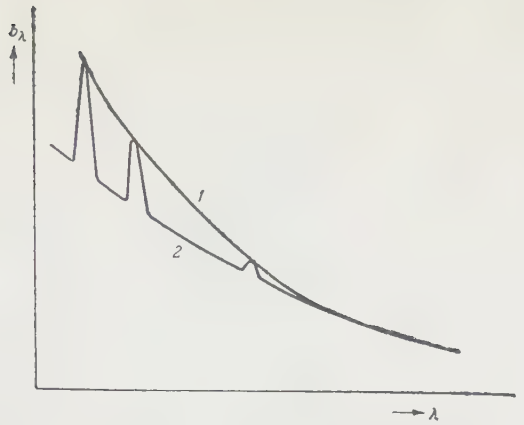


FIG. 14. Dependence of b_λ of the spark-discharge channel on the wave length for the case of total (1) and partial (2) brightness saturation.

is on the order of 9×10^4°K, and the brightness of the shock wave front, after passing through a maximum, drops to a value corresponding to an approximate temperature of 1.8×10^4°K.

It must be noted, however, that the theoretically predicted[51] brightness maximum has not yet been observed experimentally.

To check the applicability of the concepts developed by Zel'dovich and Raizer to the spark-discharge channel, it becomes very important to study the temperature distribution over the cross section of the channel. For this purpose Vanyukov and Mak[23] investigated the discharge-channel brightness for those cases which the channel plasma is partially transparent in certain regions of the spectrum. At the same time, they studied the spectral brightness density b_λ of the continuous and line radiation for varying rates of energy intake by the discharge channel. Owing to the large absorption of radiation by the plasma in the lines of the ionized atoms of the gas, maximum brightness is reached for line radiation more readily than for continuous radiation. Figure 14 shows schematically the channel-brightness spectra of a spark discharge for total brightness saturation over the entire spectrum (curve 1) and for the case when b_λ reaches a maximum value only in the line spectrum and the long-wave spectral region (curve 2).[23] The temperature, determined from the maximum value of b_λ of line-spectrum radiation, was in good agreement with the temperature determined from the maximum value of b_λ of continuous radiation. These results give grounds for assuming that the temperature distribution over the spark-discharge channel cross section does not differ greatly from a uniform distribution. Actually, taking into account the strong temperature dependence of the absorption coefficient of the plasma[52] it can be expected that were a region with higher temperature than that of the "screening" layer to exist inside the channel, the lines of the ionized atoms would be self-reversed or else would be observed in absorption.* It must be noted that G. G. Dolgov and S. L. Mandel'shtam,[55] on the basis of an examination of the gas density distribution in the spark-discharge channel, also reached the conclusion that the temperature is constant over the cross section of the channel.

The absence of considerable temperature gradients in the current-conducting discharge channel is apparently related to the considerable heat conduction (radiant and electronic) of the discharge plasma at temperatures on the order of 3 to 4 times 10^4°K.

It also follows from the results of reference 23 that the temperature of the spark-discharge channel is independent of the rate of energy intake by the channel and is approximately equal to 27,000°K, 35,000°K and 43,000°K for xenon, argon, and nitrogen respectively. The temperature values obtained for nitrogen are in good agreement with the discharge temperature determined spectroscopically from the intensity ratio of two spectral lines,[36] a method that can be applied only in the absence of noticeable reabsorption in the channel, i.e., in cases where brightness is far from being saturated.† One can therefore assume that a spark-discharge channel of constant temperature is maintained over a very wide range of variation of energy intake by the discharge channel.

The foregoing experimental results give grounds for assuming that the brightness saturation of the spark-discharge channel is related not to the screening of the high-temperature discharge zones by the cooler ones, but to the presence of a maximum channel temperature. The existance of a maximum temperature for a given gas can be explained qualitatively from energy considerations. As the channel temperature increases there is, on the one hand, an increase in the power intake due to increased ionization and radiation. On the other

*The radiation absorption spectrum of a spark discharge may show lines of neutral atoms.[54] In this case the absorption layer is the shell of the current-conducting channel, heated by the passage of the shock wave to a temperature on the order of 10^4°K. This zone absorbs very little of the continuous radiation in the visible region of the spectrum, particularly during the time of the maximum brightness.

†In reference 43 the ratio of the intensities of the lines Si IV were used to determine the channel temperature of a low-voltage spark in air (U = 200 v, C = 1000 μf, L = 10^{-6} to 10^{-3} henry). It was found to equal approximately 30,000°K. Considering that the temperature determined in this investigation was averaged over the glow time of the Si IV lines, the agreement with the results of reference 23 is satisfactory.

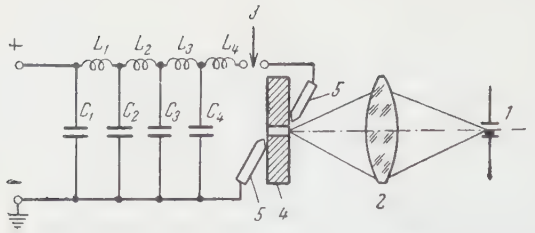


FIG. 15. Circuit diagram of a capillary spark discharge. $C_1 = C_2 = C_3 = C_4 = 100\ \mu f$, $L_1 = L_2 = L_3 = L_4 = 1.5$ microhenry. 1) slit of the spectrograph, 2) condenser, 3) spark-discharge control switch, 4) textolite plate with an opening, 5) electrodes.

hand, the increased ionization of the channel plasma lowers the resistance, causing the energy entering into the discharge channel to be reduced.[56] Effects in the same direction are obtained when the discharge-channel diameter is increased upon increase of its energy intake;[6] and also in the increase of heat conduction of the plasma at high temperatures. All this makes it impossible for the power absorbed per unit volume of the plasma channel to increase indefinitely, and this in turn limits the temperature of the channel of an open spark discharge.* It must be noted, however, that the mechanism by which the brightness becomes saturated in a spark discharge cannot yet be considered as finally established.

II. SPARK DISCHARGE IN CAPILLARIES

From the point of view of producing a high instantaneous brightness, a spark discharge with a channel is bounded by the walls of a capillary tube is of particular interest. The energy density in a so confined discharge channel can become very large, and consequently high temperatures and brightnesses can be produced in the channel. Another advantage of the capillary discharge is that the glowing substance is stable in space, an important factor in many measurements. The capillary discharge is used in optical shadowgraphic and interferometric apparatus[13,19,59] and others.

Among the first papers devoted to an investigation of high-brightness capillary spark discharges, mention should be made of that of Anderson.[25] He discharged a 1 to 2 μf capacitor at 35 kv in a tubular lamp up to 30 cm long and 0.1 to 2 cm in diameter. The tube was evacuated to pressures of the order of 1 cm Hg. The discharge was therefore produced essentially in an atmosphere of the vapor of the wall material

*From this point of view it is interesting to note the statement made in several papers[35, 57, 58] that the current density in channels of open spark discharges reaches a limiting value.

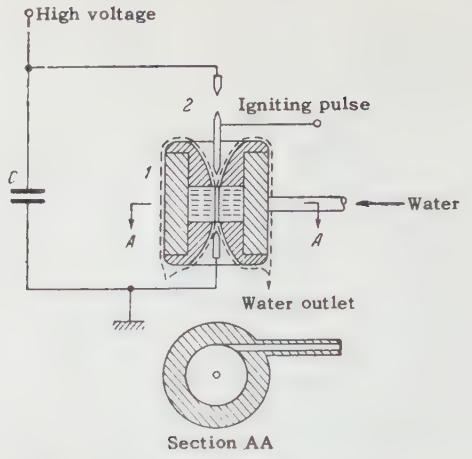


FIG. 16. Setup for producing a capillary spark discharge in a water turbine.

(silicon). The maximum temperature measured by the author amounted to 52,000°K.

Edgerton and Cathon[50] report the development of a xenon capillary pulsed tube, with channel brightness up to 10^7 c/cm^2 at flash durations from 0.3 to 3 μsec, depending on the discharge circuit parameters. The capillary is 1.2 mm in diameter and 6 mm long. The tube is made of quartz and can withstand up to a thousand flashes without a substantial reduction in the light output.

M. P. Vanyukov, A. A. Mak, and M. Ya. Ures[31] report that when a circuit with 0.011 μf capacitor charged to 29 kv is discharged into a capillary 0.4 mm in diameter, the temperature reaches 94,000°K.

N. N. Ogurtsova and I. V. Podmoshenskii[29] have developed a pulsed light source with a brightness that stays constant during the pulse (approximately 100 μsec). They used for this purpose the discharge of an artificial transmission line through an opening in a textolite plate. The arrangement is shown in Fig. 15. The capillary was 2 mm in diameter and 10 mm long. The artificial transmission line consisted of four identical sections with $C = 100\ \mu f$, $L = 1.5$ microhenry. The discharge voltage was 3 kv. The characteristic impedance of the line ($R = 0.12$ ohm) was matched to the resistance of the spark gap. The light pulse in such a discharge had an almost rectangular form, and the brightness temperature measured by the authors was 32,000°K.

A. A. Mak[32] describes a pulsed light source in which the spark discharge is produced in a cavity bounded by a water wall. The advantage of such a light source is that it can withstand very powerful discharges for an unlimited time interval without damage. A schematic diagram of the setup is shown in Fig. 16. Water enters into a turbine 1

TABLE II

Type of discharge	Gas	Gas pressure (atmos)	Discharge voltage (kv)	Capacitor rating (μf)	Capillary diameter (mm)	Temperature (10^3 °K)	Visual brightness (10^6 c/cm^2)
Capillary	Air	1	12	0.011	1.3	29	8
"	"	1	12	0.5	1.3	64	39
"	"	1	12	0.011	0.4	48	18
"	"	1	12	0.011	0.25	36	16
"	"	1	29	0.011	0.4	94	50
Open	"	1	12	0.011	—	23	7
Capillary	Xenon	4	12	0.011	2.5	23	7
Open	"	4	12	0.011	—	27	11

in a direction tangent to the wall at a pressure of 1.5 to 2 atmos. The centrifugal forces produce a cylindrical cavity along the axis of the turbine, with a diameter that is determined essentially by the diameter of the apertures in the end walls of the turbine. The air spark gap 2 serves to prevent production of an arc discharge in the capillary. The discharge circuit has a capacitance of 0.05 μf and an inductance of 0.1 microhenry. At a discharge voltage of 18 kv, the channel temperature reaches 60,000° K, corresponding to a visual brightness of 33×10^6 c/cm^2.

The optical properties of the plasma of a capillary discharge channel are investigated in references 25, 29, 31, 32, 40, 50, 60 and 61. Comparison of the channel brightness as observed from a direction parallel to the axis of the discharge and from a direction perpendicular to the axis of the discharge has shown that at temperatures higher than 40,000° K, the absorption coefficient of the channel plasma is close to unity even at thicknesses on the order of 0.05 mm.[32] That the discharge column is opaque is established also in references 25, 29, and 60. A study of the spectral distribution of the continuous radiation from a capillary discharge has shown[29,31,32] that at sufficiently large energy densities the channel radiates as a black body, whose temperature can be determined from the spectral brightness density of the channel.

Of great physical and practical interest is the question of the maximum temperatures and brightnesses that can be reached in a capillary spark-discharge channel. References 31 and 32 report on a study of the effect of the discharge circuit parameters and also of the capillary diameter on the temperature and brightness of the discharge channel. Table II lists the measurement results obtained in reference 31.

It has been established that higher temperatures and brightnesses can be obtained in a bounded spark-gap channel than in an open discharge. However, for each discharge mode there exist a certain optimum capillary diameter, at which the channel brightness has a maximum value. If the channel is either excessively or insufficiently constricted, its brightness decreases. When the energy intake by the discharge channel increases, the brightness and temperature increase but at a rate that slows down with increasing energy (Fig. 17).[32] This latter circumstance may be caused, in addition to a considerable increase in the specific heat of the gas at high temperatures, also by intense evaporation of the capillary walls in the case of powerful discharges. The influence of the wall material on the radiation from a capillary spark discharge has been investigated in references 52, 60, and 62 to 64. Weltner[64] investigated, in particular, the influence of the capillary-wall material on the intensity of the continuous radiation at a gas pressure of approximately 1 mm Hg in the capillary and at current densities of the order of 4×10^4 amp/cm^2. He established that the emission intensity is independent of the gas pressure if the amount of evaporated wall material exceeds the amount of gas that fills the capillary.

It is shown in reference 29 that at current densities on the order of 4×10^5 amp/cm^2 the radiation spectrum remains unchanged as the gas pressure is reduced from 1 atmos to 1 mm Hg. The authors of this paper have therefore reached the conclusion that the discharge originates primarily in the wall-material vapor, and that the pressure of this vapor may reach 500 atmospheres. We can thus assume that the amount of evaporated wall material increases with increasing current density

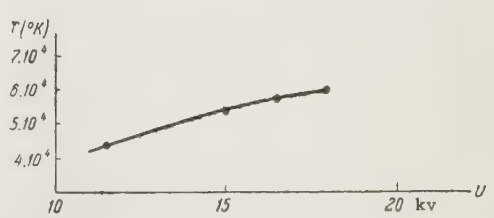


FIG. 17. Dependence of the temperature of a capillary spark discharge channel on the discharge voltage. C = 0.05 μf, L ≅ 0.1 microhenry, d = 0.5 mm.

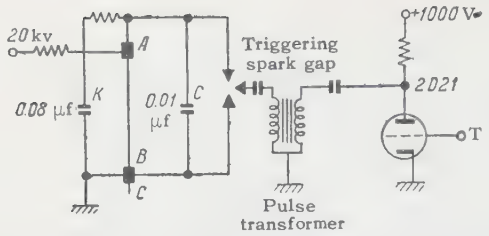


FIG. 18. Schematic wiring diagram of the guided-spark generator, the "Defatron."

in the capillary. This, on the one hand, entails a large amount of energy for evaporation, dissociation and ionization of the vapor, and on the other hand makes it difficult for the energy to enter into the discharge channel, owing to the sharp increase of pressure in the capillary. These considerations give grounds for assuming that the principal factor that limits the temperature and brightness of the spark-discharge channel in a capillary is the evaporation of the capillary walls. However, this problem must be investigated further.

III. GLIDING SPARK DISCHARGE

The spark discharge in air, the simplest pulsed light source used in high-speed photography for more than a century, yields very brief light flashes (10^{-6} to 10^{-7} seconds) but has a low light output. New light sources developed in the recent two decades, based on the use of the luminescence produced by pulsed electric discharges in heavy inert gases (argon, krypton, and xenon) have a much larger light output, but do not yield flashes shorter than several microseconds in duration. A most interesting source of light is the gliding spark discharge, in which the spark is produced between electrodes located on a surface of a dielectric, so that the discharge gap can be made 10 to 20 times longer than the gap in the ordinary spark discharge, and the light output can be substantially increased.

The gliding spark discharge has been under investigation for a long time (see references 65 to 69), but its use as a source of brief and intense light flashes is relatively recent.[13,70-74]

To facilitate the breakdown of the long spark gaps necessary to obtain a high light output, use is made of the so-called guided spark. This spark, which is used in particular in a spark generator known as the "Defatron" produces a 10^{-6}-sec light flash at a discharge energy of approximately 200 joules and a capacitor voltage of 22 kv.[13] Its operating principle can be seen in Fig. 18. The spark gap consists of two annular electrodes A and B, placed 10 cm apart on a dielectric tube, inside of which is inserted a control electrode C in the form of a rod. A potential difference of 22 kv is applied between electrodes A and B. The discharge is initiated by grounding the electrode C, which is connected through a large resistance to electrode A. The large field gradient produced between the electrodes A and C causes intense ionization of the air at the surface of the tube near electrode A, causing the gap to break down. The control electrode is grounded by means of a 3-electrode spark discharge gap triggered by applying a voltage pulse to the grid of a thyratron. The "Defatron" flash tube has a lifetime from 100 to 300 discharges.

A certain shortcoming of these sources of light, in which a gliding discharge is produced over the surface of the glass tube, is that the spark does not always follow a straight line between the anode and the cathode and may wind itself around the tube and thus partially screen the light. In addition, a light source with a straight long glowing body requires a cylindrical-parabolic reflector to concentrate the light flux, and the use of such a reflector is not always convenient. It becomes therefore advantageous to produce a light source with a more compact glowing body, suitable for use with an ordinary spherical optical condenser system. For this purpose it is necessary to be able to change the trajectory of the discharge path in the desired direction, something that can also be achieved in a guided gliding-spark discharge.

The gliding discharge breaks down along a path in which the ionization is higher than in the surrounding volume. The ionization, other conditions being equal, depends on the thickness of the dielectric between the control and principal electrodes and increases with decreasing thickness. The spark can therefore be guided by cutting a groove in the dielectric layer along the direction of the desired propagation of the discharge. Tawil[74] used synthetic rubber as dielectric and obtained S-shaped and circular discharge trajectories. Such light sources withstood up to 1200 flashes.

In another proposed illuminating device[72,73] the guided discharge is produced over the surface of a tube made of a porous ceramic, impregnated with an electrolyte, in which one end of the tube is inserted. By changing the concentration of the electrolyte it is possible to change the gap resistance. Furthermore, owing to the diffusion of the electrolyte, the characteristics of the gap are automatically restored and its service life is consequently increased. Approximately 90 percent of the energy stored in the discharge circuit is liberated in such a discharge gap. Using a 0.25-μf capacitor

charged to 12 kv, a spark-gap length of 5 to 8 cm can be reached. The half-width of the light pulse is 0.9 μsec. The intensity of the flash, as measured by its photographic action, is 5×10^6 candlepower. If the gliding discharge is produced in argon at 1 and 11 atmos, the light output increases (over that in air at atmospheric pressure) 1.5 and 4 times, respectively. However, the duration of the light flashes is also increased.

A spectroscopic investigation of the gliding spark discharge has shown[75] that in weak discharges there is a weak continuous spectrum and a line spectrum. In powerful discharges, the spectra are characterized by greatly broadened lines and by an intense continuous background, which increases with increasing pressure of the gas that surrounds the dielectric.

In conclusion it must be noted that the light characteristics of the gliding spark discharge, in spite of its advantages as a source of short and intense flashes, has not yet been fully investigated.

IV. ELECTRIC EXPLOSION OF WIRES

Intense light flashes can be obtained also when large current pulses flow through thin metallic wires. This changes the wire explosively into a metal vapor in which the electric discharge produces a high-temperature plasma.

The electric explosion of wires, first observed in 1815,[76] is still the object of research and rather lively discussions,[77-87] owing with the rather complicated mechanism by which the metal is converted from its solid state into a vapor within a very short time (10^{-6} to 10^{-7} sec).

The glow associated with the electric explosion of the wire can be used in many cases to produce intense flash illumination. Such a source of light has certain advantages over flash bulbs, for no radiation is absorbed in the walls of the bulb, and more effective use is thus made of the ultraviolet radiation. Furthermore, when exploding a wire it is possible to supply an almost unlimited amount of energy to the discharge and to obtain very high instantaneous values of light flux. Thus, for example, when photographing non-luminous phenomena through a Kerr cell, the flash produced by an electric explosion of a wire has an instantaneous light intensity of 5×10^8 candlepower.[88] An analogous source of light was used successfully in investigations involving pulsed photolysis.[89]

The temperature developed in exploding wires was determined in references 28 and 90. These measurements have shown[28] that the temperature averaged over the volume of the metal vapor cloud is 20,000 to 30,000° K at the instant of time when the expansion of the glowing vapor cloud practically ceases. The method used in reference 28 to measure the vapor temperature of the exploding wire (by determining the speed of propagation of sound waves in the vapor) does not make it possible to estimate the temperature at the early stage of the explosion, since the plasma is opaque during that time.

There is one paper especially devoted to the photometry of the light energy radiated during the explosion of the wire,[91] where the blackening of photographic plates by light from the explosion is compared with that produced by a standard source of light. A shortcoming of this technique is that it does not afford an estimate of the time variation of the radiation, and yields only the average radiation power over a certain effective flash time, the length of which is taken by the author from some other paper. In addition, the paper does not contain an estimate of the actual dimensions of the glowing body, and consequently there is no possibility of using the obtained data to determine the temperature and brightness of the glowing body.

A detailed observation of the time variation of the brightness of the discharge channel in the explosion of wires would be of great interest, since there are grounds for expecting that the metal vapor becomes heated to very high temperatures. Actually, it has been shown by calculation,[81] on the basis of results of direct oscillographic determination of the energy liberated during the explosion, that in the case of the explosion of a silver wire 0.15 mm in diameter and 3 cm long mounted in organic glass, the energy liberated through the discharge of a 10 μf capacitor charged to 7 kv is sufficient to heat the vapor to 150,000° K (disregarding the radiation and ionization losses).

V. SHOCK WAVES

The propagation of strong shock waves in gas is accompanied by an intense glow, the mechanism of which is described, for example, in the thorough survey of Ya. B. Zel'dovich and Yu. P. Raizer.[53]

At the present time the glow of shock waves is used as a pulsed source of light in special explosion lamps, usually filled with argon.[92-97]

The arrangement of one such lamp has been described in reference 96. The lamp is a glass bulb 6.25 cm in diameter, filled with argon. The lamp contains a charge of explosive substance of conical form, with a base diameter somewhat smaller than the diameter of the bulb. The distance between the base of the charge and the end window of the lamp varies from 0.5 to 2 cm and determines the duration of the flash. The lamp

TABLE III

Gas	Temperature, °K		
	Spark discharge	Shock wave	
	Experimental value[23]	Experimental value[24]	Theoretical value[24]
Argon	35 000	34 000	60 000
Krypton	—	33 000	90 000
Xenon	27 000	30 000	106 000

operates in the following manner. An electric detonator mounted at the vertex of the conical charge initiates the detonation of the explosive, and at the instant when the detonation wave reaches the base of the charge a shock wave is excited in the argon and produces an intense glow. The duration of the resultant flash is determined by the time required for the shock wave to cover the distance from the surface of the charge to the surface of the end window of the bulb.

Figure 19 shows the time variation of the light flashes in explosion pulsed lamps filled with argon, for various distances between the charge and the glass. The flash of such a lamp has an instantaneous light intensity of 225 million candlepower at a glow area of 33 cm^2 corresponding to an average brightness of 6.8×10^6 c/cm^2. These data show that the explosion lamps produce very brief and intense light flashes.

I. Sh. Model'[24] measured directly the brightness of the front of strong shock waves propagating in an atmosphere of air or heavy inert gases. His method, photographic photometry of time-scan pictures of the shock waves, made it possible to determine the front temperature from its brightness and to measure simultaneously the propagation speed of the wave front. It is possible to calculate from this speed the temperature that can be attained in shock compression. It follows from theoretical considerations that the highest temperature, and consequently the highest brightness, can be expected on a front of strong shock waves propagating in heavy inert gases.

A comparison of shock-wave temperatures calculated from hydrodynamic theory with those determined from the brightness of the front leads to interesting conclusions. The calculated dependence of the temperature of the shock-wave front on the speed of its propagation, taking into account the energy lost to ionization and thermal radiation, is shown in Fig. 20 along with experimental data for argon, krypton, and xenon. The great discrepancy between the theoretical and experimental values of the temperature cannot be attributed to measurement errors, which do not exceed ±20 percent whereas, for example in the case of xenon, the experimental value of the temperature is 3.5 times smaller than the theoretical one. It must also be noted that in contrast to theoretical expectations, the experimental temperature maximum is observed in argon, and the temperature decreases consecutively from argon to krypton to xenon. In his treatment of this interesting phenomenon I. Sh. Model' follows the point of view developed by Ya. B. Zel'dovich and Yu. P. Raizer.[51-53]

Interesting conclusions can be drawn from a comparison of the maximum temperatures produced in heavy inert gases by the propagation of shock waves and by a spark discharge. The values of these temperatures, taken from references 23 and 24, are listed in Table III.

A striking fact is the close agreement between

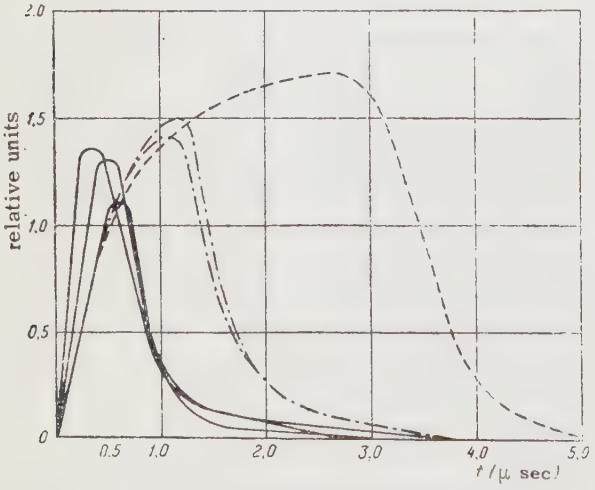


FIG. 19. Dependence of intensity of radiation on the time for flashes of argon explosion lamps. —— Argon layer 0.5 cm thick; –·–·– argon layer 1.0 cm thick; ---- argon layer 2.0 cm thick.

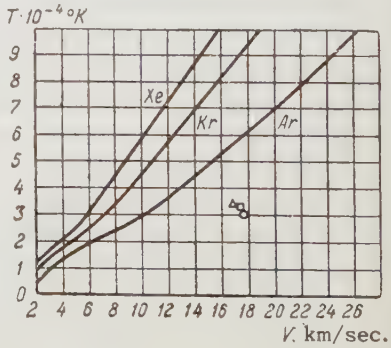


FIG. 20. Temperature of the front of a strong shock wave in inert gases. Experimental values: Δ) argon, □) krypton, ○) xenon.

TABLE IV

T °K	$2\cdot10^3$	$3\cdot10^3$	$4\cdot10^3$	$5\cdot10^3$	$6\cdot10^3$	10^4	$2\cdot10^4$
B c/cm²	44,3	$2,83\cdot10^3$	$2,34\cdot10^4$	$8,41\cdot10^4$	$2\cdot10^5$	$1,1\cdot10^6$	$6,2\cdot10^6$

T °K	$3\cdot10^4$	$4\cdot10^4$	$5\cdot10^4$	$6\cdot10^4$	$1\cdot10^5$	$2\cdot10^5$
B c/cm²	$1,21\cdot10^7$	$1,82\cdot10^7$	$2,45\cdot10^7$	$3,1\cdot10^7$	$5,63\cdot10^7$	$1,23\cdot10^8$

the numerical values of the temperatures measured in the spark-discharge channel and in shock waves. Furthermore, in either case the maximum temperature is in argon and the minimum temperature is in xenon. All this suggests that the plasma produced by electric current has much in common with plasma produced by shock compression, and, apparently, the maximum temperature is the same in either case.

VI. POSSIBILITY OF FURTHER INCREASE IN THE BRIGHTNESS OF PULSED LIGHT SOURCES

The data obtained in recent years on the light-emission characteristics of pulsed light sources permit certain predictions on the possibility of a further increase in the brightness of such light sources.

It must be noted beforehand that the problem of increasing the brightness is essentially a problem of raising the temperature. Table IV gives the dependence of the visual brightness B of a black body on its temperature.

As can be seen from the table, as the temperature of the black body increases the brightness first increases very strongly, and then the increase slows down until, at temperatures above 60,000°K, there is an almost linear relation between the visual brightness and the temperature. Naturally, in addition to having a high temperature, the source of radiation must have also a sufficiently great optical depth in order to produce a high brightness.

The results obtained in references 15, 22, 23, and 26 have shown that maximum brightness that can be obtained in a spark-discharge channel increases with diminishing atomic weight of the gas in which the discharge occurs.*

Consequently, from the point of view of increasing the brightness of pulsed light sources, it becomes of considerable interest to study further the spark discharge in such gases as neon and helium.[15,26] It is also of interest to investigate the temperature and brightness in the spark-discharge channel in other gases, particularly in hydrogen at high pressure.[9]

It must be borne in mind here that the higher the maximum brightness obtainable in a gas, the more difficult it is to produce the discharge conditions under which this maximum brightness is obtainable. Consequently, further improvement in discharge circuits is required, through reduction of the inductance and increase in the discharge voltage. Certain other difficulties must also be overcome in the development of pulsed lamps filled with neon and helium at high pressure.

High temperature and brightness can also be produced in a spark-discharge channel by raising the emitted-energy density. Of interest in this respect is the capillary spark discharge[31,32,50] and the use of the magnetic field of the discharge current to confine the discharge channel.

As regards the capillary spark discharge, further research is needed to increase the brightness of this type of discharge. It is also necessary to study the effect of the wall material of the capillary on the temperature of the discharge channel, and to improve the operating characteristics of capillary light sources.

Limitation of the spark-discharge channel by the magnetic field of its own current (pinch effect), observed in gases at high pressure, is reported in references 44 and 98. Under these conditions one can expect to produce very high temperatures and brightnesses in the spark-discharge channel. However, this problem has hardly been studied hitherto.* Nor have all the potentialities been investigated for increasing the temperature and brightness in the case of shock waves. In particular, it becomes interesting to study the brightness originated in reflection of a shock wave from a barrier, and in

*There are grounds for assuming that this is related to the increased ionization potential and the investigated gas, not to the atomic weight. In polyatomic gases (N_2, O_2) the dissociation energy apparently also becomes important.

*V. S. Komel'kov and D. S. Parfenov[44] give for the channel temperature a value of 33 electron volts (T = 250,000°K) for a discharge in air of atmospheric pressure.

collision between shock waves.[41,53,98,99] According to theoretical estimates, temperatures on the order of 10^9 °K should be obtainable in these cases.[100]

[1] R. Post, Revs. Modern Phys. **28**, 338 (1956);

[2] W. Finkelnburg, Kontinuierliche Spektren, Berlin, 1938.

[3] A. Unsöld, Ann. Physik **33**, 607 (1938).

[4] G. Glaser, Optik **7**, 33 (1950).

[5] H. F. Talbot, Phil. Mag. **3**, 73 (1852).

[6] M. M. Gegechkori, J. Exptl. Theoret. Phys. (U.S.S.R.) **21**, 493 (1951).

[7] K. S. Vul'fson and I. Sh. Libin, J. Exptl. Theoret. Phys. (U.S.S.R.) **21**, 510 (1951).

[8] Vanyukov, Dobretsov, Isaenko, and Mak, Светотехника (Illumin. Eng.) No. 4, p. 9 (1958).

[9] H. Fischer, J. Opt. Soc. Amer. **43**, 394 (1953).

[10] C. V. Boys, Nature **47**, 415 (1893).

[11] F. Früngel, Optik **3**, 128 (1948).

[12] L. S. G. Kovasznay, Rev. Sci. Instr. **20**, 696 (1949).

[13] P. Fayolle and P. Naslin, J. Soc. Mot. Pict. Engrs. **60**, 603 (1953).

[14] F. Früngel, Z. angew. Phys. **6**, 183 (1954).

[15] H. Fischer, J. Opt. Soc. Amer. **47**, 981 (1957).

[16] H. Fischer, Physik. Verhandlung. **6**, 177 (1955).

[17] J. D. Craggs and J. M. Meek, Proc. Roy. Soc. **186**, A, 241 (1946).

[18] J. W. Beams, A. R. Bulthau, A. C. Lepsley, J. H. McQueen, L. B. Shoddy and W. D. Whitehead, J. Opt. Soc. Amer. **37**, 868 (1947).

[19] J. A. Fitzpatrick, J. C. Hubbard and W. J. Thaler, J. Appl. Phys. **21**, 1268 (1950).

[20] V. S. Komel'kov and G. N. Aretov, Dokl. Akad. Nauk SSSR **110**, 559 (1956).

[21] S. Magun and S. Woerner, Z. angew. Phys. **10**, 41 (1958).

[22] Vanyukov, Mak, and Parazinskaya, Оптика и спектроскопия (Optics and Spectroscopy) **1**, 642 (1956).

[23] M. P. Vanyukov and A. A. Mak, Dokl. Akad. Nauk SSSR (1958) (in press).

[24] I. Sh. Model', J. Exptl. Theoret. Phys. (U.S.S.R.) **32**, 714 (1957), Soviet Phys. JETP **5**, 589 (1957).

[25] J. A. Anderson, Astrophys. J. **75**, 394 (1932).

[26] F. A. Charnaya, Оптика и спектроскопия (Optics and Spectroscopy) **4**, 725 (1958).

[27] H. E. Petschek, P. H. Rose, H. S. Glick, A. Kane and A. Kantrowitz, J. Appl. Phys. **26**, 83 (1955).

[28] J. A. Anderson and S. Smith, Astrophys. J. **64**, 295 (1926).

[29] N. N. Ogurtseva and I. V. Podmoshenskii, Оптика и спектроскопия (Optics and Spectroscopy) **4**, 539 (1958).

[30] Vanyukov, Isaenko, and Khazov, J. Tech. Phys. (U.S.S.R.) **25**, 1248 (1955).

[31] Vanyukov, Mak, and Ures, Оптика и спектроскопия (Optics and Spectroscopy) **4**, 90 (1958).

[32] A. A. Mak, Dokl. Akad. Nauk SSSR (1958) (in press).

[33] W. Weizel and R. Rompe, Theorie elektrischer Lichtbögen and Funken, Leipzig, 1949.

[34] W. Weizel and R. Rompe, Ann. Physik **1**, 285 (1947).

[35] N. N. Sobolev, J. Exptl. Theoret. Phys. (U.S.S.R.) **13**, 137 (1943).

[36] S. L. Mandel'shatam and N. K. Sukhodrev, J. Exptl. Theoret. Phys. (U.S.S.R.) **24**, 701 (1953).

[37] S. L. Mandel'shtam and N. K. Sukhodrev, Izv. Akad. Nauk SSSR, Ser. Fiz. **19**, 11 (1955) [Columbia Tech. Transl. **19**, 7 (1955)].

[38] Vaĭnshteĭn, Leontovich, Malyavkin, and Mandel'shtam, J. Exptl. Theoret. Phys. (U.S.S.R.) **24**, 326 (1953).

[39] S. L. Mandel'shtam and I. P. Tindo, Izv Akad. Nauk SSSR, Ser. Fiz. **19**, 60 (1955) [Columbia Techn. Transl. **19**, 57 (1955)].

[40] A. A. Babushkin, J. Exptl. Theoret. Phys. **14**, 279 (1944).

[41] G. Glaser, Z. Naturforsch. **6a**, 706 (1951).

[42] L. Huldt, Spectrochim. Acta **7**, 264 (1955).

[43] J. van Calker and H. Braunisch, Z. Naturforsch. **11a**, 612 (1956).

[44] V. S. Komel'kov and D. S. Parfenov, Dokl. Akad. Nauk SSSR **111**, 1215 (1956), Soviet Phys. "Doklady" **1**, 769 (1956).

[45] L. Sptizer, Physics of Fully Ionized Gases (Russ. Transl.) IL, Moscow 1957.

[46] Vul'fson, Libin, and Charnaya, Izv. Akad. Nauk SSSR, Ser. Fiz. **19**, 61 (1955).

[47] I. S. Abramson and N. M. Gegechkori, J. Exptl. Theoret. Phys. (U.S.S.R.) **21**, 484 (1951).

[48] F. A. Charnaya, Оптика и спектроскопия (Optics and Spectroscopy) **1**, 857 (1956).

[49] H. Fischer, Tele-Tech and Electronic Industries, No. 5, 15 (1956).

[50] H. Edgerton and P. Cathon, Rev. Sci. Instr. **27**, 821 (1956).

[51] Ya. B. Zel'dovich, J. Exptl. Theoret. Phys. (U.S.S.R.) **32**, 1126 (1957), Soviet Phys. JETP **5**, 919 (1957).

[52] Yu. P. Raĭzer, J. Exptl. Theoret. Phys. (U.S.S.R.) **33**, 101 (1957), Soviet Phys. JETP **6**, 77 (1958).

[53] Ya. B. Zel'dovich and Yu. P. Raĭzer, Usp. Fiz. Nauk **63** (1957).

[54] M. P. Vanyukov and L. D. Khazov, Dokl. Akad. Nauk SSSR **42**, 523 (1953).

[55] G. G. Dolgov and S. L. Mandel'shtam, J. Exptl. Theoret. Phys. (U.S.S.R.) **24**, 691 (1953).

[56] G. Glaser and D. Sautter, Z. Physik **143**, 44 (1955).

[57] I. S. Abramson and I. S. Marshak, J. Tech. Phys. (U.S.S.R.) **12**, 632 (1942).

[58] I. S. Marshak, J. Exptl. Theoret. Phys. (U.S.S.R.) **16**, 703 (1946).

[59] M. P. Vanyukov, J. Tech. Phys. (U.S.S.R.) **16**, 889 (1946).

[60] A. A. Babushkin, J. Exptl. Theoret. Phys. (U.S.S.R.) **14**, 156 (1944).

[61] A. A. Babushkin, J. Exptl. Theoret. Phys. (U.S.S.R.) **14**, 184 (1944).

[62] O. Hahn and W. Finkelnburg, Z. Physik **122**, 36 (1944).

[63] H. Greiner, Naturwiss **40**, 238 (1953).

[64] K. Weltner, Z. Physik **136**, 631 (1954).

[65] G. Mierdel, Handbuch der Experimentalphysik **13**, v. 3, 282, 1929, Leipzig.

[66] F. Merrill and A. Hippel, J. Appl. Phys. **10**, 873 (1939).

[67] F. Merrill and A. Hippel, J. Appl. Phys. **21**, 1269 (1950).

[68] E. Hueter and H. Pappen, Elektrotechn. Z. **74**, 15 (1953).

[69] C. Meyer, Umschau **55**, 175 (1955).

[70] E. Fünfer, Z. angew. Phys. **1**, 295 (1949).

[71] H. Schardin and E. Fünfer, Z. angew. Phys. **4**, 185 (1952); H. Schardin and E. Fünfer, Z. angew. Phys. **4**, 224 (1952).

[72] H. Luy and R. Schade, Z. angew. Phys. **6**, 253 (1954).

[73] H. Luy and R. Schade, Actes 2-ème Congres intern. photogr. et cinematogr. ultra-rapides, Paris, Dunod., 1956.

[74] E. P. Tawil, Proc. Third. Intern. Congress on High-Speed Photography, London 1957.

[75] R. Aumout and B. Vodar, Actes du 2-ème Congres intern. photogr. et cinemotogr. ultra-rapides, Paris. Dunod., 1956.

[76] G. Singer and A. Crosse, Phil. Mag. **46**, 161 (1815).

[77] N. N. Sobolev, J. Exptl. Theoret. Phys. (U.S.S.R.) **17**, 986 (1947).

[78] S. V. Lebedev and S. É. Khaikin, J. Exptl. Theoret. Phys. (U.S.S.R.) **26**, 629 (1954).

[79] S. V. Lebedev and S. É. Khaikin, J. Exptl. Theoret. Phys. (U.S.S.R.) **26**, 721 (1954).

[80] W. H. Conn, Z. angew. Phys. **7**, 539 (1955).

[81] Kvartskhava, Plyutto, Chernov, and Bondarenko, J. Exptl. Theoret. Phys. (U.S.S.R.) **30**, 42 (1956), Soviet Phys. JETP **3**, 40 (1956).

[82] I. F. Kvartskhava, J. Exptl. Theoret. Phys. (U.S.S.R.) **30**, 621 (1956), Soviet Phys. JETP **3**, 787 (1956).

[83] Kvartskhava, Bondarenko, Meladze, and Sulidze, J. Exptl. Theoret. Phys. (U.S.S.R.) **31**, 737 (1956), Soviet Phys. JETP **4**, 637 (1957).

[84] S. V. Lebedev, J. Exptl. Theoret. Phys. (U.S.S.R.) **32**, 199 (1957), Soviet Phys. JETP **5**, 243 (1957).

[85] W. Müller, Z. Physik **149**, 397 (1957).

[86] W. M. Conn, Naturwiss. **45**, 6 (1958).

[87] E. Fünfer, M. Keilhacker and G. Lehner, Z. angew. Phys. **10**, 157 (1958).

[88] Heine-Geldern, Pugh and Foner, Phys. Rev. **79**, 230 (1950).

[89] G. K. Oster and R. H. Marcus, J. Chem. Phys. **27**, 189 (1957).

[90] M. Vaudet, Ann. de Phys. **9**, 645 (1938).

[91] W. M. Conn, J. Opt. Soc. Amer. **41**, 445 (1951).

[92] A. Michel-Levy A. and H. Muraour, C. R. Acad. sci. **204**, 576 (1937).

[93] H. Muraour, Chimie et industrie **47**, 3 (1942).

[94] H. Muraour, A. Michel-Levy and E. Vassy, Rev. optique **20**, 161 (1942).

[95] P. M. Fye, J. Soc. Mot. Pict. Engrs. **55**, 414 (1950).

[96] C. H. Winning and H. E. Edgerton, J. Soc. Mot. Pict. Engrs. **59**, 178 (1952).

[97] Sewell, Cosner, Wedaa and Gallup, J. Soc. Mot. Pict. Engrs. **66**, 21 (1957).

[98] J. A. Allen and J. D. Craggs, Brit. J. Appl. Phys. **5**, 446 (1954).

[99] O. Preining, Usp. Fiz. Nauk **55**, 595 (1955); [Oester, Chem. Z. **55**, 5/6 (March, 1954)].

[100] E. Sänger, Z. Naturforsch. **6a**, 302 (1951).

Translated by L. E. Bergstein and J. G. Adashko

Meetings and Conferences

PHYSICAL METHODS OF INVESTIGATION OF THE STRUCTURE OF MOLECULES

E. I. FEDIN

Usp. Fiz. Nauk **66**, 131-139 (October, 1958)

Expanded session of the Scientific Council for "Theory of Chemical Structure, Kinetics, and Reactivity" in the Division of Chemical Sciences, Academy of Sciences, U.S.S.R.

THE Scientific Council for "Theory of Chemical Structure, Kinetics, and Reactivity" of the Division of Chemical Sciences of the Academy of Sciences, U.S.S.R. convened on May 14 and 15, 1958 at the Institute of Chemical Physics, Academy of Sciences, U.S.S.R. to hold an expanded session devoted to the status of physical experimental methods for the investigation of the structure of molecules. The session attracted great interest and more than 300 persons participated.

In his opening address, Academician V. N. Kondrat'ev formulated briefly the scope of the broadened session, the presentation of a clear picture of the status of different physical methods for the investigation of the structure of molecules at the scientific institutions of the Soviet Union, the adoption of specific measures for further expansion of the scale of application of these methods, and the disclosure of fields where scientific progress is still noticeably lagging.

The status of diffraction methods in the research on the structure of molecules (x-ray diffraction, electron diffraction, neutron diffraction) was discussed by Prof. **A. I. Kitaigorodskii**, who divided the diffraction methods into three large groups.

The first group deals with stereochemical problems in the determination of the spatial arrangement of the atoms in the molecule. Here the interatomic distances can be measured with an accuracy to 0.1 A, while the valence angles can be determined accurate to several degrees. Detailed research in stereochemistry has recently uncovered new approaches towards the understanding of the reactivity of the molecule. Diffraction methods can render great help in this field.

The second group involves the study of the relative placement of molecules in crystals. This is the principal method of investigating intermolecular interactions. The accuracy required here is also not too high; x-ray diffraction and neutron diffraction solve such problems with sufficiently good results.

The third group of problems deals with the determination of the interatomic distances and valence angles in molecules with such a degree of accuracy, that the results can be compared with those obtained by theoretical calculations. The prevalent opinion is that this problem is basic to diffraction methods and that principal attention should be paid to it. However, the method is at the limit of its capabilities when used for solving these problems. It is necessary to take into account the crudeness of the theory, the difficulty of introducing corrections for absorption in the sample, and the fact that diffraction methods yield a structure that is an average over all the elementary cells, regular and distorted. In organic chemistry, a substance containing 0.1% of impurities is considered to be quite pure. In such a crystal, however, one encounters one molecule of impurity for each nine molecules along any direction. This impurity distorts the lattice in a volume that exceeds considerably the volume per single molecule. This is why a precise determination of the structure of the molecule can be recommended only for substances of exceptional purity and exceptional crystal quality, i.e., this problem cannot be considered principal in diffraction methods.

When estimating the relative capabilities of electron, x-ray, and neutron diffraction methods, it must be pointed out that the exposures used in these methods are related roughly speaking as $10^{-6}:1:10^{6}$. The exceedingly long exposures required prevent an extensive use of neutron diffraction. This method can, however, yield important results, for it makes it possible to determine the coordinates of the atoms of light elements down to hydrogen. Electron diffraction also has certain advantages over x-ray diffraction in the determination of the coordinates of the hydrogen atoms, but is greatly inferior to the latter in accuracy. The accuracy of diffraction methods is proportional to the number of reflexes. A good structural result is one in which from 1,000 to 2,000 diffraction dots are used; an electron diffraction pattern rarely contains more than 200. There are

many additional complications. Nevertheless, the electron diffraction method can be a valuable adjunct to x-ray diffraction studies of substances in the solid state. A structural investigation of liquids does not promise any interesting results. A most promising trend is electron diffraction of gases in those cases, when it is necessary to determine experimentally two or three parameters while the remaining parameters can be assumed to be accurately known. The number of problems in which the investigator is interested in the molecule only is quite considerable. Electron diffraction of gases deserves therefore our serious attention.

Progress in diffraction methods is quite slow in our country, owing to a shortage of laboratory centers. It is also difficult to obtain certain types of equipment. In conclusion, A. I. Kitaigorodskii introduced several proposals aimed at further advances in diffraction methods for the study of the structure of molecules.

B. S. Neporent started his talk on electron spectroscopy of molecules by dividing the problems in this field into two large groups: electron spectroscopy of simple molecules, and a study of complex molecules that have continuous spectra.

This subdivision is connected with the substantial difference between the methods of investigation. The study of a developed vibrational spectrum of a molecule, consisting of well resolved lines is a purely spectroscopic problem. The low-selectivity spectra of the complex molecules is a more difficult problem, in which it is impossible to confine oneself to obtaining the spectral characteristics. It is necessary to resort to luminescent investigations and to investigations of the state of polarization, of the glow of the molecules, of the action of the solvent, of the effect of complex formation, etc. Such investigations are carried out quite extensively in the U.S.S.R. and occupy a prominent position in the world's science. For many systems, reliable results have already been obtained on various molecular characteristics of interest to chemists: the mobility of the electron shell, the probability of the thermal activity upon variation of the electron state, the production of the biradical state, etc. A study has been made, for example, of the suppression of the photochemical processes in many synthetic materials based on organic compounds, a factor that lengthens the service life of such materials.

The Institute of Chemical Physics of the U.S.S.R. Academy of Sciences is engaged in studies on the spectra of radicals, aimed at determining the role of radicals in chemical reactions. However, we lag foreign science in this very promising field as far as the scale of the research goes.

High-level research on spectra of organic substances in the crystalline state is being done under the leadership of A. F. Prikhod'ko. It has been possible to obtain interesting data on the reactivities of molecules.

Research on complex molecules in the gas phase is being carried out on a small scale (Moscow State University and Leningrad State University). The spectra of adsorbed molecules are investigated, and the changes in the spectra yield information on the energy properties of the adsorption phenomena.

The most research is being done on polyluminescence — optical investigations involving luminescence of complex molecules and resulting in spectra that have no vibrational structure. This work is done essentially by Soviet scientists who use original methods to study the transfer of excitation energy from one molecule to another and to measure the concentration of the investigated molecules in the solid, liquid, or gaseous state.

In conclusion, B. S. Neporent noted that the status of spectroscopic apparatus, from the point of view of assortment and quality of instruments produced, is satisfactory, but the volume of production is frequently inadequate.

E. V. Shpol'skii devoted his lecture to the spectroscopy of polyatomic organic molecules. The lecturer defines "spectroscopy" as the production and measurement of sharp spectra, from which wavelengths can be determined with an accuracy to not less than 0.1 A. If such a spectrum is obtained, it is possible to determine from it the energy levels, to ascertain the symmetry of various states, to determine the transition probability, etc.

Until recently, more attention has been paid to absorption spectra than to emission spectra. Yet emission spectra, obtained by optical excitation, i.e., luminescence spectra, help solve many important problems. Such spectra of organic molecules disclose not only the allowed transitions, but also forbidden ones and help study the probabilities of many energy processes in the molecules. It is customary to assume that to obtain sharp spectra it is necessary to investigate the substance in gaseous state. In the case of polyatomic organic molecules the situation is just the opposite: gas spectra of such molecules are quite smeared out, owing to the abundance of natural oscillations and the presence of a quasi-continuum of the rotational states in the polyatomic molecules. Consequently, it is necessary to reduce the temperature, and to strive as close as possible to absolute zero, i.e., to proceed to the observation of crystal spectra. Such investigations are carried abroad out at hydrogen

and helium temperatures, and it is proposed to begin the study of spectra at helium temperatures in this country in the nearest future. When using single crystals, this procedure permits the use of polarized light, thus greatly facilitating the interpretation of the results.

However, spectroscopy of molecular crystals has two very serious shortcomings. First, the interaction of the molecules in the crystal lattice causes the appearance of the so-called lattice lines. Second, in complex molecules such as condensed aromatic hydrocarbons (pyrene, coronine, etc.), the spectra are smeared out even at helium temperatures, owing to the strong resonance effects (the sharing of electrons). The California group found a way out by preparing a substitution solution of the investigated substance in an isomorphous crystal solvent, so that the investigated molecules could be fixed at sufficient distances from each other. It was necessary in this case that the crystal solvent be transparent in the region where the spectrum of the dissolved substance was located. This occurs very rarely (McClure found only three pairs of substances that give solid solutions satisfying the above requirement).

E. V. Shpol'skii used the freezing of weak solutions of the investigated hydrocarbons in normal paraffins (pentane, hexane, and higher). In this case the luminescence spectra (singular and triplet) of the polynuclear atomic hydrocarbons, in pure substances in crystalline state (polycrystalline and single-crystal), which produce even at low temperatures a single wide continuum extending over several thousands of reciprocal centimeters, and split up even at nitrogen temperature (77°K) into a multitude of narrow lines. At the hydrogen temperature (20°K) there occurs a further narrowing of the lines and new fine details appear. The sharpness of the lines is such that it is possible to undertake a vibration analysis of the spectra (coronene, pyrene, 3, 4-benzpyrene) and to carry out in many cases a qualitative analysis of natural mixtures and artificial products. Whatever absorption spectra could be obtained appeared in the first absorption region, like the luminescence spectra.

M. M. Sushchinskii lectured on "Methods of Vibrational Spectroscopy." Thirty years have passed since the discovery of Raman scattering, and infrared spectra have been investigated for 50 years. Infrared spectra of nearly 20,000 substances have been obtained, and Raman spectra are available for more than 2,000 substances.

The list of products is continuously expanding: organic and inorganic substances, various commercial products, natural substances, and various biological objects. Studies are made on gases, liquids, solids, glasses, and polymers. The lecturer pointed out the inadequate level of most researches: the investigators confine themselves to plotting the spectrum without interpretation and without attempts to obtain from their curves the physical characteristics of the molecules. This is frequently explained by the fact that the regularly available instruments, ISP-51 and IKS-11, do not permit a detailed study of vibrational spectra.

M. M. Sushchinskii briefly described the long known capabilities of the methods of vibrational spectroscopy (determination of number of definite structural groups in the molecule, determination of interatomic forces, explanation of the electron-optical properties of the molecule), and then listed the following recent new research trends:

(1) Establishment of the presence of rotary isomers, a helpful procedure in the investigation of high-polymer compounds.

(2) Spectroscopy of compressed gases.

(3) Study of infrared spectra in polarized light, which sometimes permits location of the bonds in polymers and biological objects.

(4) Investigation of the rotational structure of infrared vibrational spectra and purely rotational Raman spectra, which makes it possible to determine the moments of inertia of molecules and hence the interatomic distances and valence angles.

(5) Raman spectra, taken near the region of the natural electron absorption of the molecule, yield information on the electrostatic properties of the molecules (P. P. Shorygin).

(6) Infrared spectra of molecules adsorbed on the surface (A. N. Terenin) clarify the nature of adsorbing centers.

(7) Owing to the tremendous volume of experimental material, the documentation of vibrational spectra assumes great importance.

Comparison and foreign works shows that the ideas and the skill of the leading Soviet investigators are the equal of those abroad, but that the working conditions are much worse here than, for example, in the U. S. The shortage of help, the inadequacy of the technical equipment, and the great lack of interest on the part of theoretical physicists in molecular spectroscopy are to blame.

The paper by **V. I. Divanov-Klokov** was devoted to infrared spectroscopy of compressed gases. When a pure gas or a mixture of gases are compressed, the interaction between a given molecule and its neighboring one can lead to a change in the symmetry, causing in turn a removal of the forbiddenness of certain transitions (forced absorption).

It was established that the intensity of absorption corresponding to vibrational frequencies of hydrogen, nitrogen, or oxygen increases in proportion to the square of the pressure, indicating an interaction between two molecules.

In 1955 Kittelar observed in gas mixtures absorption frequencies corresponding to differences and sums of all the vibration frequencies of the corresponding molecules. This was treated as a simultaneous vibrational transition in two molecules. It turned out to be more convenient technically to carry out these experiments in liquid mixtures. This was followed by an experimental verification of earlier assumptions on the possibility of simultaneous electron transitions in two colliding oxygen molecules.

The lecturer noted that little material was accumulated during the ten years elapsed since the beginning of these investigations. This was attributed to great difficulties with apparatus, since the observation turned out to be possible in the most favorable case (mixture of a polar and a non-polar gas) only at pressures on the order of 20 atmos. To observe the effect in a mixture of non-polar gases it is necessary to have hundreds of atmospheres, and in the case of electron transitions in oxygen, the necessary pressure amounts to thousands of atmospheres.

The spectroscopy of compressed gases can yield information on intermolecular interactions, particularly on the formation of unstable complexes. Considerable progress has been made in the theory of forced absorption. No theory of simultaneous transitions has been developed so far. Experiments in this field have been successful in the U.S.S.R. only at the INEOS, (Institute of Elementary Organic Compounds) Academy of Sciences, U.S.S.R. (I. V. Obreimov).

Further progress in the research is greatly hindered by the lack of apparatus capable of plotting a spectrogram within one or two minutes, by difficulties in producing sufficiently low temperatures, and by the lack of high-pressure cuvettes (up to 3500 atmos).

Lively discussions followed the first five lectures.

M. A. El'yashevich supported A. I. Kitaigorodskii's proposal to organize research on specific groups of substances by different methods. Such investigations are frequently hindered by the lack of the necessary pure substances. A radical solution to the problem of research at hydrogen and helium temperatures must be found. New electronic computers are necessary for theoretical work in the field of molecular structure.

Comrade **Makarov** discussed the application of electron beams to the study of excitation levels of molecules. Such investigations call for apparatus with mechanical workmanship of very high grade. It is necessary to organize production of suitable equipment in the plants.

Comrade **Zamkov** called for further research on the application of the Kerr effect. This procedure makes it possible to obtain rapidly accurate data for the determination of the structure of molecules and for analytical purposes. No suitable apparatus has yet been produced, and should be manufactured as soon as possible.

Comrade **Levin** spoke of the unhappy state in the manufacture of modern infrared spectrometers. An impression is gained that plants are afraid of complicated electronic circuits. The proposed organization of a card file of molecular spectra should be rapidly put into practice.

E. V. Shpol'skii noted the obvious unfavorable situation in theoretical research in the field of electronic and electron-vibrational spectra. The attention devoted by theoretical workers to nuclear physics is understandable and justified, but this does not mean that it is necessary to refrain from work on solid-state physics. So far there are no centers for systematic training of young theoreticians, whose interests would lean towards these difficult problems. Semiqualitative model representations are very important to experimenters, and of great help in their work. It is also necessary to develop classical computational methods with use of the most modern computational techniques. The most important problem is to bridge the gap between us and the U. S. with respect to the number of cryogenic installations. It is necessary to publish an atlas of absorption spectra.

N. D. Sokolov also noted that among the 70 or 80 theoreticians, who graduate every year from the Physics Faculty of the Moscow State University, only one or two remain to work in molecular physics. Certain changes are necessary in the training of theoretical physicists, so that the students do not get the idea that molecular physics is an "uninteresting" field. N. D. Sokolov supported the proposal of A. I. Kitaigorodskii that a new journal, "The Structure of Matter," be published.

B. S. Neporent acquainted the participants of the session with the work of the Commission on Spectroscopy on documentation of molecular spectra. Foreign catalogues should be acquired in the near future to accelerate this work. In spite of the general high level of Soviet spectroscopic research, we rarely publish work in which experimental mastery is combined with so high a theoretical level.

It is necessary to organize theoretical groups in laboratories. To improve the apparatus situation substantially, it is necessary to broaden the instrument-building base.

Ya. L. Gol'dfarb recalled that for five or ten years principal attention was concentrated in chemistry on the nature of the bond. What is of primary importance now is the dependence of the reactivity on the stereochemistry of the molecules in the ground and in the transition states. In the development of physical methods of help to chemists, it is therefore necessary to remember above all the stereochemical problems. Interchange of experience must be organized so that workers in various institutes can turn to others where some method is being developed for advice or for material help. It is necessary to welcome initiative shown by the Scientific Council in demonstrating to a large circle of persons the various tools that physics places at the disposal of chemistry.

G. L. Slonimskii has emphasized that all modern methods acquire particular significance in the investigation of the structure of polymers. It is here that the greatest number of physical measurements are carried out and that the least are reported. It is advisable to devote a special session of Scientific Council to the results of the investigation and the structure of high-molecular substances by various physical methods.

A. A. Brandt delivered a lecture on the electric methods of investigating the structure, properties, and interaction of molecules. Measurement of the dielectric constant makes it possible to determine the value of the dipole moment, the period of dipole oscillations, the relaxation time, etc. Measurement of the electric losses makes it possible to measure small concentrations of polar groups, inaccessible to infrared spectroscopy. Losses in crystals are determined by the density of packing, i.e., experiment in this field give information on the interaction and structure of molecules. This uncovers a way towards solving problems of importance to radio engineering, such as the production of new materials with low losses. Unfortunately, existing methods of measuring losses are still insufficiently accurate. To increase the significance of electric methods of research on the structure of molecules it is necessary to have closer contact between physicists and chemists. It is necessary to train radiophysicists in the measurement of different properties of matter. It is necessary to produce standard apparatus for the measurement of dielectric constants, to publish a collection of materials on the dielectric properties of matter, and to organize a laboratory for research on the electric properties of matter.

Comrade **Ivanov** supported the proposed development of apparatus for the measurement of electric characteristics of matter in a wide range of frequencies and temperatures.

The session of May 15 opened with a lecture by **A. M. Prokhorov,** devoted to gas radiospectroscopy and electronic paramagnetic resonance. The advantage of gas radiospectroscopy is that one can work with small quantities of matter: a small cavity need be filled with gas at a pressure of 10^{-2} mm mercury. Absorption takes place at centimeter waves. The object of study are simple molecules with electric and magnetic dipole moments. Complicated molecules give a large number of overlapping weak lines. Isotope substitutions shift the lines. From the intensity is is possible to estimate the value of the potential barrier of the torsional vibration about a single bond. A recent development is a theory whereby these barriers are determined from frequency measurements. If the molecule contains nuclei with non-vanishing electric quadrupole moments, the splitting of the lines of the rotational spectrum make it possible to determine the constant of the quadrupole bond of the free molecule. These data can be compared to measurements of the quadrupole bond in the solid state, carried out by the method of the nuclear quadrupole resonance. The lecturer has noted that the number of problems solved by gas radiospectroscopy is not very extensive. It is therefore meaningful to make use of this method only if a specific research plan has been formulated. The lecturer then described the capabilities of electronic paramagnetic resonance as regards the study of the structure of molecules and the kinetics of chemical reactions. It becomes possible to clarify the valence state of paramagnetic ions and to observe transition from one valence to the other. In liquids it is possible to study the solvate shells and processes of complex formation. It is easy to observe a small number of free radicals, occurring in chemical reactions. The significance of research in the free radicals for chemistry cannot be overestimated, for here is a literally unchartered territory as far as work is concerned. This is why electronic paramagnetic resonance is gaining more and more recognition. The essential problem here is the acquisition of modern apparatus and good magnets.

L. A. Blyumenfel'd devoted his paper to applications of electronic paramagnetic resonance in biology and in the study of methylorganic compounds. In the laboratory of V. V. Voevodskii (Institute of Chemical Physics, Academy of Sciences, U.S.S.R.) an investigation of organic compounds containing strontium and chromium has led to far-reaching conclusions concerning the electron structure of

these compounds. A recently-observed interesting phenomenon is that the hyperfine structure of paramagnetic-resonance lines depends on the nature of the solvent and on small additives of various substances to the solvent. The theory of this phenomenon has not been developed but is apparently will prove useful for research on the possibility of formation of solvate complexes etc.

Another intersting trend in the work of this laboratory is research in radiation chemistry — the study of free radicals formed by gamma irradiation of hydrocarbons and high polymers.

Irradiation of frozen hydrocarbons produces a large yield of alkyl free radicals, the structure of which is determined uniquely from the hyperfine structure of the absorption lines. It is also possible to observe in these experiments signals of free atomic hydrogen. Both the free radicals and the atomary hydrogen can be stored in frozen state for many days. When the temperature is increased it is possible to observe the recombination kinetics.

Electron paramagnetic resonance is used in L. A. Blyumenfel'd's laboratory to study the mechanism of fermentative catalysis. Definite conclusions have been drawn concerning the structure of albumen ferments: it has been ascertained that an excess electron obtained in one manner or another is delocalized to a considerable extent over the albumen molecule. This makes it possible to tackle the problem of changes produced in albumens by muscular contractions etc. In conclusion, L. A. Blyumenfel'd noted the need for regular production of spectrometers for the study of electronic paramagnetic resonance.

A report by **B. M. Kozyrev** on the work done at the Laboratory for Electronic Paramagnetic Resonance in Kazan' was read by N. D. Sokolov. Information was obtained in this laboratory on the symmetry of local electric fields that act on the ion from the solvate shell, on the lower boundary of the lifetime of the solvates, and on the partially-covalent character of the bonds between the central ion and the molecules of the solvate shells. The character of the chemical bonds between paramagnetic ions and their surroundings can be studied also in glass-like systems. An investigation is being made of the influence of adsorption of paramagnetic gases (O_2 and NO) on the width of the absorption line in diphenyl pycril hydrazil at various temperatures. Above 0°C, the adsorbed oxygen forms on the surface of the free radical a so-to-speak two dimensional liquid; when the temperature is reduced, the motion of the adsorbed O_2 molecules "freezes." In addition, work is being done on proton resonance in carbons and other carbon-containing substances and apparatus for observing the quadrupole resonance is under study. At the Kazan' State University and the Kazan' Branch of the U.S.S.R. Academy of Sciences, extensive theoretical and experimental research is being done on spin-lattice relaxation in salt crystals and in electrolyte solutions.

N. D. Sokolov noted the high level of the work of the Kazan' group, where a genuine collaboration between theoreticians and experimenters has been established. There are still many unsolved problems in the theory of radiospectroscopy. Among the most important ones is that of quadrupole interactions, theoretically very difficult, which can yield important data on the character of bonds, on the distribution of the electron density, etc.

N. D. Sokolov then took the floor to deliver a paper on the use of nuclear magnetic resonance in chemical research. If a constant external magnetic field H_0 is imposed on a specimen, containing nuclei with non-zero magnetic moment μ, this magnetic moment can become oriented in different manners relative to the external field, each orientation corresponding to a definite energy level. The distances between the neighboring levels are $\mu H_0/I$, where I is the nuclear spin. If the constant magnetic field is modulated by a weak radio-frequency field at a frequency given by $h\nu = \mu(\mu/I)H_0$, the orientation changes, i.e., resonance absorption of the energy of the radio-frequency field takes place. Strictly speaking, the field in which the nucleus is located is not equal to H_0. It is less than H_0 by a value σH_0, due to the influence of the electron shell. The constant of magnetic screening, σ, is sensitive to variations in the number of electrons surrounding the given nucleus. For example, protons having different electron surroundings give somewhat different resonant frequencies, and this manifests itself in a fine splitting of the line of nuclear magnetic resonance. In addition to this fine splitting (chemical shifts), one can observe in the magnetic spectra a hyperfine multiplet structure, due to the interaction between paramagnetic nuclei of different atoms with each other through the electron spins. A characteristic quantity is also the spin-lattice relaxation time, determine from the rate of change of the signals under definite conditions, and also the spin-spin relaxation time, determined from the widths of the signals.

Measurement of the chemical shifts makes it possible to determine the structure of the molecules. Thus, a molecule of ethyl alcohol contains three different protons: the proton of the CH_3, group, the proton of the CH_2 group, and the proton

of the OH group. The nuclear magnetic resonance spectrum of ethyl alcohol, when properly resolved, therefore consists of three lines, whose intensities are related as 3 : 2 : 1. An analysis of the hyperfine multiplet structure discloses not merely the presence of similar groups, but also their location. Research of this type is carried out in our country at present at the Leningrad University under the leadership of F. I. Skripov and at the Moscow University under Yu. S. Konstantinov, using flourine containing compounds. Research on chemical shifts in proton resonance is being done extensively abroad.

A deeper study of the structure of the molecule by this method requires a theoretical analysis of the connection between the values of the chemical shifts and the structural molecule. Only the first steps have been made in this direction thus far.

Interesting results were obtained in the study of the hydrogen bond by means of proton magnetic resonance. The speed of the reaction can be estimated from the form of the signal. A. I. Rivkind uses the sensitivity of the spin-lattice relaxation time to the presence of paramagnetic impurities to determine the character of the interaction of the solvate shell with the central ion.

It must be noted that the level of Soviet research in the field of nuclear magnetic resonance lags considerably the parallel work abroad. This is explained by the lack of apparatus.

The character of the requirements imposed on nuclear magnetic spectrometers suitable for the study of the structure of molecules is detailed in the paper by **N. M. Ievskaya.** It is necessary to distinguish between broad lines and narrow lines of nuclear magnetic resoance. Broad lines have widths from 0.1 to 10 gausses. To record such lines, the apparatus must discern from 0.01 gauss to 1 gauss. Since experiment is carried out in fields on the order of 10 gausses, the resolution of the apparatus should be in this case from 10^{-4} to 10^{-6}. This is a requirement imposed on the homogeneity of the field, on its stability, and on the stability of the frequency of the field. When working with broad lines the signals obtained are weak and it therefore becomes necessary to take use more material, on the order of 10 cm^3. In this case 1 cm^3 should contain from 10^{16} to 10^{18} nuclei of the investigated element. When studying the chemical shifts it is necessary to deal with lines ranging from 0.1 to 100 milligausses in width. The resolution required in this case is from 10^{-6} to 10^{-9}. It is impossible to obtain such great homogeneity in large volumes; the specimens must be on the order of 0.1 cm^3. In these experiments 1 cm^3 of the specimen should contain 10^{17} to 10^{18} nuclei of the investigated element. The chemical shifts increase with increasing mass number. The chemical shifts for hydrogen nuclei do not exceed 10 milligausses, i.e., the spectrometer resolution must be better than 10^{-7}. A resolution of 10^{-5} to 10^{-6} is sufficient to observe the chemical shifts of fluorine. This is exactly the resolution of our best apparatus "home made" by the researchers themselves in Leningrad, Moscow, and Kazan'. Yet in the U. S. the Varian company produces commercial nuclear magnetic spectrometers, which have excellent characteristics. Our lag in this field will not be overcome until the laboratories are provided with modern equipment. In the conclusion of her lecture N. M. Ievskaya made several specific recommendations for intensification of work in nuclear magnetic resonance.

V. L. Tal'roze stated certain ideas on the organization of regular production of high-grade radiospectroscopic apparatus.

E. I. Fedin remarked that a new method of investigating molecular and crystalline structure of solids — nuclear quadrupole resonance — did not yet acquire in our country a status commensurate with its significance and with the potentialities of the method. Nuclear quadrupole resonance is unique as regards to sensitivity to chemical effects. While chemical shifts in nuclear magnetic resonance give only the fine structure of the line, requiring resolution to 10^{-8}, similar effects shift the frequency of the quadrupole resonance by several megacycles, the frequency itself being on the order of 30 Mcs. Crystal effects cause splitting of the absorption lines of tens and hundreds of kcs at line widths on the order of 1 to 3 kcs. In the laboratory of A. I. Kitaigorodskii (Institute of Elementary Organic Compounds, Academy of Sciences, U.S.S.R.) a quadrupole radio spectrometer was placed in operation during the last year and projects are under way in this field. The first favorable results were obtained recently in Leningrad and Kazan'. Nuclear quadrupole resonance eliminates the difficulties involved with stabilization of the magnetic field, but the effect is weak. It is necessary to cover a broad band of frequencies (from 1 Mcs to 1,000 Mcs) and to measure the resonance frequencies accurately. Progress is slowed by the lack of many instruments, the acquisition of which ahould be arranged for.

K. V. Vladimirskii spoke about his investigations, which uncovered the possibility of research on chemical shifts not only in hydrogen and fluorine, but in all nuclei with magnetic moments. Usually experiments on chemical shifts are performed under conditions close to stationary, when it is im-

possible to attain the sensitivity necessary for the observation of all nuclei. In a specially selected non-stationary mode (deep modulation of the magnetic field of a relatively high frequency) it becomes possible to obtain spectra which are just as easy to read as those obtained with slow coverage of the resonance region. It turns out that the requirements for apparatus are less stringent in this case. One can measure chemical shifts, for example, in silicon-29 etc., a very important factor in chemistry.

M. A. El'yashevich expressed a few opinions on progress in radiospectroscopic methods. It is necessary to solve quickly the problem of providing suitable apparatus, to join forces of the interested workers, and put an end to home-made equipment. The most important theoretical problem is that of quadrupole interactions, a problem involving complicated mathematical computations of multi-electron systems.

E. V. Shpol'skii indicated that many fine effects, studied by magnetic-resonance methods can be studied with sufficient reliability by using the fine structures of the electron spectra plotted at low temperatures. It has been established that the occurrence of intramolecular hydrogen bonds removes the fine structure. The same effect is obtained by the formation of a glass-like state. Naturally, it is necessary to have diffraction gratings of high resolution, but they are easier to prepare than radiospectrometers.

L. L. Dekabrun spoke of the many advantages of the pulse technique of observation of nuclear resonance ("spin-echo"). If two successive radio frequency pulses of suitably selected frequency and spacing are applied to a specimen, a "spin-echo," which is a pulsed nuclear-resonance signal is produced in the receiving equipment. This relaxes substantially the requirements concerning the homogeneity of the magnetic field, and the apparatus becomes simple and reliable. The procedure is indispensable if the relaxation times play the principal role, and also in research of exploratory character. Steps must be taken towards the application of pulse methods to nuclear resonance.

Ya. G. Dorfman delivered a paper on magnetochemistry. Strictly speaking magnetic methods of molecular research should include not only the measurement of susceptibility but also all varieties of magnetic-resonance experiments. Unfortunately radiospectroscopic investigations are usually separated from purely magnetic ones.

Statistical methods of measuring the susceptibility are applicable to liquids and solids. Measurements of susceptibilities of gases and vapors are still in the development stage. For such a simple object as methane, various authors give experimental values that differ by a factor of more than three. Recently radiotechnical methods have been developed for the measurement of susceptibility, and these increase substantially the speed of the experiment. Apparently this method can be used to measure the susceptibility of not only the end products of the reaction, but also of the products produced furing the time of reaction.

Magnetic susceptibility can usually be represented in the form $\chi = \chi_{p0} + \chi_d + \chi_p$, where χ_{p0} is the orientational paramagnetic susceptibility, χ_d the diamagnetic susceptibility (the Langevin precession), χ_p the paramagnetic term connected with the deformation of the electron shells. χ_{p0} is connected directly with the measurement of the number of unpaired spins; the value of χ_d must be taken into account to determine χ_{p0}. The diamagnetic correction for light molecules is small, but in many cases it cannot be neglected. A way out can be found by combining measurements of χ with the use of paramagnetic resonance. At a sufficient intensity of the radio-frequency field it is possible to "turn off" the orientation term and to measure only χ_d, thus estimating the dimensions of the orbit. Furthermore, when investigating mixtures of substances it is possible to "turn off" individual components of the mixture. No such work has been performed anywhere. The lecturer reports on the results of his latest investigations: he has found out that χ_d exhibits a pronounced dependence on the electric polarizability, i.e., χ_d can be calculated for a large number of molecules with great accuracy. This makes it possible to determine χ_p, something no one has engaged in. Knowledge of χ_p will make it possible to estimate the asymmetry of the electron shell. All this is outwardly reminiscent of Pascal's scheme, but the Pascal experiments has no physical meaning while the proposed additive scheme is based on modern concepts of the mechanism of the phenomenon. Magnetochemistry can deepen substantially the results of research on the electronic structure of free radicals, obtained with the aid of paramagnetic resonance. A very promising study is that of magnetic properties of molecules during the course of chemical reactions. In conclusion, Ya. G. Dorfman noted the quite unsatisfactory situation with respect to magnetochemical research in the Soviet Union and proposed certain measures to improve the situation.

N. D. Sokolov agreed with the poor opinion of the state of our magnetochemistry and supported

Ya. G. Dorfman's proposals. N. D. Sokolov praised highly the establishment of a linear connection between the magnetic susceptibility and the electric polarizability.

M. A. El'yashevich again called for close contact between different fields of research and for a rapid elimination of the lag in some of the neglected sections of our science, such as magnetochemistry. Another area is radiospectroscopy of molecular beams.

A paper on mass-spectrometric research on the structure of molecules was read by **V. L. Tal'roze.** Although mass spectrometry gives less data on the structure than optics or radiospectroscopy, the information it does yield is sometimes unique. The principal data that can be obtained are the ionization potential of the molecule and the binding energy of the molecule. Ionization is effected by electron impact in deep vacuum, and is followed by mass spectroscopic analysis of the resultant ions. The ionization potentials can be estimated in this case accurately only if the electron beam is sufficiently monochromatic. This problem has not been fully solved, but it becomes possible to obtain satisfactory accuracy. Recently the accuracy has been increased through the use of photoionization. The binding energy is more difficult to determine, owing to the indeterminacy in the states of the electron-impact products. It becomes necessary to complicate the research substantially by carrying out a series of measurements and by leaning on thermochemical data. In addition to the binding energy and the ionization potential, it becomes possible to obtain qualitative information on the structure of molecules by splitting them with a beam of electrons having energies on the order of several thousands electron volts and by subsequent analysis of the fragments. Here we can speak only of the positions of the double bonds, and of the positions of the large atom groups that are bound to carbon or to other heavy atoms. The position of the hydrogen and of several other atoms cannot be established, since strong regroupings occur inside the excited ion during such an ionization. Investigations of this type, cannot be carried out with the mass spectroscopes produced by our industry. It would be necessary to organize production of additional adapters for the regular-production instruments, permitting rapid conversion of the mass spectroscopes to the measurement of ionization potentials and binding energies. As to the analytic applications of mass spectroscopy, the situation in this field is satisfactory from the point of view of quality. We lag behind the U. S. or England in the number of projects and in the use of mass spectroscopy in manufacture.

In his concluding remarks, Academician **V. N. Kondrat'ev** noted that several years ago our lag in research on the structure of molecules was considerably greater than now, and that the situation has been substantially improved. Many investigators, whose interests have been hitherto confined to apparatus, begin to be interested in molecular structure once they start operating the apparatus. There is no doubt that when more equipment becomes available the lag in some fields will be rapidly overcome.

On May 22 the scientific council adopted an extensive resolution concerning the proceedings of its broadened session of May 14-15.

Translated by J. G. Adashko

From the Current Literature

DIRECT OBSERVATION OF OPTICAL VIBRATIONS OF A CRYSTAL LATTICE OF A SOLID BY MEANS OF NEUTRON SCATTERING

P. CHENTSOV

Usp. Fiz. Nauk **66**, 347-348 (October, 1958)

THE modern view[1] is that the vibrations of a crystal lattice can be represented in the form of a set of "normal vibrations," each corresponding to an elastic wave propagating in the direction of the wave vector **k** (the absolute value of which is $2\pi/\lambda$, where λ is the wavelength) at a frequency $\nu = \omega/2\pi$. An important characteristic of any particular lattice is its vibration spectrum, or the dependence of ω on k. This dependence, even in the case of the simplest crystals, is given by several curves of two distinct types, the so-called acoustic and optical modes. Physically, the difference between these two types is that in the former case the near atoms vibrate "in phase," like neighboring volume elements of a continuous elastic medium, while in optical oscillations the neighboring atoms vibrate "in anti-phase" and move "opposite" each other. The $\omega(k)$ curves of the optical mode lie above the acoustic curves (i.e., they are characterized by a higher frequency). They differ also in that ω depends less on k, that ω is finite when $k = 0$ (in the case of acoustic vibrations $\omega = 0$ when $k = 0$), and that ω is not greatly influenced by the direction of **k** in the crystal.

Experiments on scattering of slow neutrons offer a powerful means of investigating the vibration spectrum of solids. The theory of scattering of neutrons by crystals, developed in its day by I. Ya. Pomeranchuk[2] and Wick,[3] has received further development in recent years[4] and has served as the basis for numerous experimental investigations.[5] In particular, these investigations have clearly confirmed the character predicted for the spectrum by the modern theory of lattice dynamics, namely the absence of a sharp discontinuity in the $\omega(k)$ curve (the Debye theory of specific heat is known to call for such a discontinuity existed at certain values of ω and k), and the presence of a double maximum on the $\omega(k)$ curve for certain lattices such as vanadium (in the acoustic mode). Two recent papers[6,7] report that by scattering of cold neutrons it is possible to observe directly, for the first time, optical lattice vibrations in two substances, single-crystal germanium and polycrystalline zirconium hydride. The same experimental procedure was reported in both papers. A broad neutron beam from a nuclear reactor was passed through a filter made of polycrystalline beryllium. Neutrons with a de Broglie wavelength less than 3.96 A experienced Bragg reflection from the corresponding beryllium lattice planes and were scattered, while the slower neutrons passed through the filter without great attenuation. This produced intense beams of slow neutrons with little divergence and with an average energy of approximately 0.004 ev. This beam was incident on the investigated target. A mechanical selector with rotating slits was used to sort neutrons scattered at given angles by their time of flight. The selector used in these experiments had a sufficiently high resolving power to permit observation of sharp peaks in the energy distribution of the scattered neutrons.

It was found that the scattered neutrons have considerable energies (compared with the energies of the incident neutrons), amounting to hundredths and tenths of an electron volt. This indicates that the scattered neutron acquires some energy from the vibrating lattice (the experiments were performed at room temperature and at higher temperatures). In the case of zirconium hydride, for example, this leads[6] to a continuous energy distribution (Fig. 1) with a broad maximum near 0.25 ev. This corresponds to energy transferred to the neutrons by lattice vibrations of the acoustic type. However, as can be seen from an examination of Fig. 1, a sharp peak is clearly pronounced against the background of this distribution near 0.134 ev

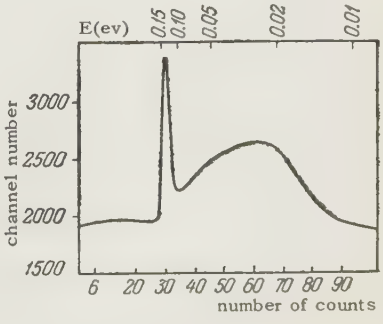


FIG. 1

(±0.015 ev). This peak corresponds to optical lattice vibrations with energy 0.130 ev. A similar energy peak, corresponding to optical vibrations, is observed[6] also in neutron scattering by a single crystal of germanium. True, the optical peak is accompanied in this case by several other peaks, corresponding to coherent scattering of neutrons by acoustic vibrations. This optical peak can nevertheless be readily separated. It is located at a higher energy (0.038 ev). Furthermore, the energy of this peak (unlike those of other peaks) remains constant as the germanium crystal orientation is varied in the instrument. Since the scattered neutrons interact at various orientations with lattice vibrations corresponding to different values of k, the fact that the peak energy is independent of the orientation indicates that the energy and frequency of the corresponding lattice vibration are independent of k, which is precisely the characteristic feature of optical vibrations.

In the case of zirconium hydride, it was possible to investigate the observed optical vibrations in greater detail.[7] It can be assumed that in this case we deal with vibrations of the hydrogen atom in an isotropic potential well of parabolic shape. If this is so, there should exist a whole series of optical levels with energies $E_n = nh\nu$ (where $n = 1, 2, 3 \ldots$). Indeed, the raising of the temperature of a zirconium-hydride target to 393°C resulted in a weaker (and more diffused) peak, $E_2 \approx 2h\nu$, in the spectrum of the scattered neutrons along with the above-mentioned principal peak ($E_1 = h\nu$). It has been found that the population of the first optical level (a measure of which is obviously the intensity of the principal peak in the spectrum of the scattered neutrons) varies with temperature in accordance with the Boltzmann law $\exp(-E_1/\kappa T)$, with $E_1 = -0.130 \pm 0.005$ ev. This value coincides with the directly-determined value of the energy of the optical vibrations.[6]

A recent report[8] states that scattering of slow neutrons has made it possible to determine directly, for the first time, the energy spectrum of elementary excitations in liquid helium II. The reviewed papers demonstrate the new opportunities which this progressive method offers to researchers in the field of solid state physics.

---

[1] See, for example, M. Born and K. Huang, Dynamical Theory of Crystal Lattices (Russ. Transl.) IL, Moscow, 1958 (Oxford, 1954).

[2] I. Ia. Pomeranchuk, Phys. Z. Sowjetunion **13**, 65 (1938).

[3] G. C. Wick, Physik. Z. **38**, 403, 689 (1937).

[4] G. Placzek and L. van Hove, Phys. Rev. **93**, 1207 (1954) et al.

[5] B. N. Brockhouse, Trans. Roy. Soc. Can. III, 49 (1955), paper 118. B. Jacrot, Compt. rend. **240**, 745 (1955). Carter, Hughes, and Palevsky, Phys. Rev. **104**, 271 (1956). B. N. Brookhouse and A. T. Stewart, Revs. Modern Phys. **30**, 236 (1958). A. T. Stewart and B. N. Brookhouse, Revs. Modern Phys. **30**, 250 (1958). Eisenhouer, Pelah, Hughes, and Palevsky, Phys. Rev. **109**, 1046 (1958), et al.

[6] Pelah, Eisenhouer, Hughes, and Palevsky, Phys. Rev. **108**, 1091 (1957).

[7] Anderson, McReynolds, Nelkin, Rosenbluth, and Whittemore, Phys. Rev. **108**, 1092 (1957).

[8] See Usp. Fiz. Nauk **65**, 545 (1958).

Translated by J. G. Adashko

Personalia

ANATOLII BOLESLAVOVICH MLODZEEVSKII

(on his seventy-fifth birthday)

S. N. RZHEEVKIN

Usp. Fiz. Nauk **66**, 145-146 (September, 1958)

A. B. Mlodzeevskii celebrated his 75th birthday on April 11, 1958. His father was Professor B. K. Mlodzeevskii of the Moscow University, a great mathematician famous as a lecturer. An uncle of his was also a professor at the Moscow University.

Mlodzeevskii began his scientific work while still a student, in the laboratory of the renowned physicist P. N. Lebedev. His first scientific work, a highly refined experiment, led to the proof that the velocity of sound is constant in air at all frequencies up to ultrasonic.

During the 55 years of his scientific and pedagogical activity, Mlodzeevskii made many original scientific investigations in various branches of physics.

Mlodzeevskii's principal field of scientific research was geometric thermodynamics, i.e., the thermodynamic basis of one of the most important branches in physico-chemical analysis, namely the study of the diagrams of state of systems in equilibrium. This study began with Gibbs's researches at the end of the 19th century.

Mlodzeevskii adapted creatively the ideas of the originator of the theory of the equilibrium of heterogeneous systems (Gibbs) and of the founder of physical-chemical analysis (N. S. Kurnakov). His extensive original and profound research on the most important problems of geometric thermodynamics has greatly enriched this science.

Mlodzeevskii made an exhaustive analysis of the basic problem of the shape of the thermodynamic potential curve of a binary system in which a chemical compound is formed. This analysis resulted in his theory of the singular points on the melting curve. This theory provided a thermodynamic foundation for the experiments of N. S. Kurnakov and his school. Particularly important is Mlodzeevskii's investigation of the shapes of the melting curves of chemical compounds that form solid solutions with their components. One of these very interesting investigations concerned the thermodynamic surfaces of single-component systems, in which he established the connection between the four characteristic functions — the total energy, the free energy, the enthalpy, and the thermodynamic potential — using a brilliant application of "contact transformations."

After encountering in his experimental work an exceedingly complicated and previously unobserved phenomenon, namely the existence of liquid crystals with two dissociation temperatures (upper and lower) in the ketyl alcohol-chloresterin system, Mlodzeevskii succeeded in finding a correct interpretation of these phenomena through his deep insight into the theory of phase equilibrium.

Mlodzeevskii has been engaged for several decades in extensive fruitful pedagogical work in higher institutions of learning. His profound scientific erudition and inherited talent as a brilliant lecturer place him in the forefront among the professors of the Moscow University. He always delivers his

lecture material clearly and concisely and accompanies it with brilliant demonstrations. His lectures are always very popular among the students. During his career at the University Mlodzeevskii devised more than 200 new demonstrations in physics, of which more than 100 were original.

During his many years at the University and other higher institutions of learning, Mlodzeevskii lectured on general physics, electrodynamics, thermodynamics, physics of metals, phase theory, crystal optics, vector analysis, procedures of physics, and others. He wrote many texts of university level, including "Molecular Physics," "Thermodynamics," and "Phase Theory." Particularly noteworthy is the college text he edited for instructors, "Lecture Demonstrations in Physics," in eight volumes, four of which were written by him personally.

Mlodzeevskii's scientific-social activities began even before the October revolution. He participated actively in the organization of the publicly-endowed library of the Moscow Physical Society. This library played a very important role during the period when the Moscow physicists, headed by P. N. Lebedev, struck at the Moscow University in protest against the reactionary politics of the Tsarist Minister Kasso. During the first imperial war, Mlodzeevskii actively organized x-ray divisions in hospitals. After the October revolution, he engaged constantly in various social activities at the Physics Faculty and in the House of Scientists. He delivers annual lecture-demonstrations for students and teachers in secondary schools.

A. B. Mlodzeevskii is a man of high culture and inexhaustible energy. In him we have a foremost scientist, working without fatigue in the field of urgent scientific problems, and a brilliant pedagogue and professor, giving more and more of his knowledge and experience to the younger generation.

For the many years of selfless work for the benefit of the Fatherland, Mlodzeevskii was awarded the Order of Lenin and medal "for valor in work."

On April 23 a celebration was held at the physics faculty of the Moscow University in honor of the famous hero of the day. Mlodzeevskii gave a summary report of his activities and showed many beautiful lecture demonstrations. The Ministry of Higher Education, the Administration of the Moscow State University, and the Administration of the Physics Faculty congratulated the guest of honor and thanked him for his many years of fruitful scientific-pedagogical activity. The attending representatives of many organizations, his coworkers, and his friends warmly congratulated Mlodzeevskii and wished him vigor and health, that he continue to follow with his customary energy the fruitful scientific and pedagogical activity for the benefit of our great Fatherland.

Translated by J. G. Adashko

BOOK REVIEWS

MIKHAIL ANDREEVICH SHATELEN

Bibliographic index. Compiled by A. I. Isachenko and K. I. Shafranskii. Edited by V. S. Ravdonik, Academy of Sciences (U.S.S.R.) Press, 1958, 198 pages (Library of the Academy of Sciences, U.S.S.R. and of the M. I. Kalinin Leningrad Polytechnic Institute). Price, 3.20 rub. Printing, 1700 copies.

The life and activity of Mikhail Andreevich Shatelen (1866-1957) was linked most intimately with the development of electrical engineering in our country. He began his scientific and engineering activity in the late 1880's, when only the first timid steps were made in Russia towards the commercial use of electricity. Throughout his lifetime, Shatelen was a great social worker — the chairman of the Electrotechnical Division of the Russian Technical Society, an active participant in numerous electrotechnical congresses, conferences, and conventions in our country and abroad, and chairman of the board of the All-Union Scientific Engineering Technical Society for Power and Electric Communication (VNITOE). His part in the development of the plan for the electrification of the Soviet Union (GOELRO), drafted at the initiative of V. I. Lenin, is well known.

A brief bibliography of the published works of M. A. Shatelen, listing 123 titles, appeared in N. A. Shost'in's "Mikhail Andreevich Shatelen on his 80th Birthday," Moscow-Leningrad, 1946. A much more complete "Bibliography of the Works of M. A. Shatelen and of the Literature Concerning Him," compiled by the basic library of the M. I. Kalinin Leningrad Polytechnic Institute and containing 370 titles, was published in Leningrad in 1956. The library of the U.S.S.R. Academy of Sciences and the M. I. Kalinin Leningrad Polytechnic Institute has now published an incomparably more detailed and more extensive annotated bibliography, in the form of a separate book (compiled by A. I. Isachenko and K. I. Shafranskii).

The book begins with a review of Shatelen's life and activity written by V. S. Ravdonic, chairman of the Department of General Electrical Engineering of the Leningrad Polytechnic Institute. It is accompanied by a 22-page detailed list of the principal dates in his life and activity.

The main portion of this book contains an index to Shatelen's books and articles and detailed annotations. The index is arranged in chronological order and contains 400 titles (including translations by Shatelen and works published under his editorship). Shatelen was a direct active participant in the development of electrical engineering in prerevolutionary Russia and in the U.S.S.R., and came in contact with most prominent Russian inventors and electricians, constantly contributing to the promotion and practical realization of their discoveries and intentions. All of Shatelen's notes on the history of electrical engineering are exceedingly valuable. An undoubted advantage of the annotations is therefore the fact that their compilers note particularly all reference to the history of electrical engineering contained in Shatelen's writings.

Shatelen was a pioneer in the teaching of electrical engineering in our higher institutions of learning. His many lithographed lecture notes were written with great mastery and originality. Unfortunately, these lithographed editions were not preserved in their entirety even by our largest libraries. It is therefore important that the titles of Shatelen's lecture notes (which were repeatedly reprinted) are exhaustively listed in the new index of his works, as are the articles written for the student newspapers of the Polytechnic Institute, such as "Tovarishch", "Industrial'nyi," "Politekhnik" (items 380 to 420 in the bibliography).

The bibliographic index concludes with a detailed list of the literature on Shatelen. In this list of 118 titles there are mostly articles and notes by his contemporaries and students published in journals and partially in newspapers.

Many years have elapsed since the start of Shatelen's activity. It is therefore necessary to approach with increasing caution the recording of facts that have occurred 70 years ago. Unfortunately, this caution has not always been observed in the list of dates in Shatelen's life and activity. Thus, on p. 37 it is stated that in 1888 Shatelen was "awarded the degree of Candidate of Physico-Mathematical Sciences," whereas at that time the Candidate degree was granted by the university without reference to any specialty. In the same page it is stated that Shatelen "translated into Russian many works on electromechanics," although his translations dealt more with electrical engineering as a whole (see items 352 and 354 in the bibliography). On the next page, 38, it is mentioned that in 1891 Shatelen became the "laborant" of Professor O. D. Khvol' son, without explaining that the duties of "laborant" corresponded at that time to the present duties of an assistant. Since 1892 Shatelen contributed to the journal "Elektrichestvo" and "later on became the secretary of the editorial board and a member of the editor-

ial committee" (p. 38); it should also be added here that he was once editor of the journal. In 1893 Shatelen was "invited to join Electrotechnical Institute" (page 35); the word "invited" should be more accurately replaced by the words "passed the competitive examination."

It is stated on page 39 that, in 1900 the International Congress of Electricians in Paris elected Shatelen "vice president of the section on electric measurements." Actually he was the vice president of the entire congress. It is noted on page 56 that in "1946 to 1954 Shatelen participated in the work of the Commission on the History of Physico-Mathematical sciences of the Academy of Sciences, U.S.S.R." Yet he was invited to join this commission by Academician A. S. Krylov earlier, in 1945.

Item 30 of p. 65 lists an obituary of N. G. Slavyanov written by Shatelen. It should have been noted that this obituary was based on the personal recollections of Shatelen. On page 164 the initials of I. G. Klyapkin are erroneously given.

Another remark concerns the styling of the publication. The book would undoubtedly gain in appearance had the titles of the cited references been printed in bold-face type.

Apart from the above remarks, which refer merely to details, the new bibliography of Shatelen's work makes a very favorable impression. The list is very complete and will be used for reference not only by biographers of Shatelen but also by all those engaged in the history of electrical engineering in our country.

— M. Radovskii

Usp. Fiz. Nauk **66**, 147-148 (September, 1958)

NEW BOOKS IN PHYSICS

*Aglintsev, K. K. Дозиметрия ионизирующих излучений (Dosimetry of Ionizing Radiations) Second, revised edition. Moscow, Gostekhizdat, 1957, 503 pp., illustr., bibliography pp. 492-499, 7,000 copies, 18.50 rub.

Babakov, I. M. Теория колебаний (Theory of Oscillations). Textbook for higher technical institutions of learning. Moscow, Gostekhizdat, 1958, 628 pp., illustr. 15,000 copies, 11.90 rub.

Contents: Part I. Linear Systems with Finite Degrees of Freedom. Chapters: 1. Systems with One Degree of Freedom. 2. Equations of Small Oscillations of a System with Several Degrees of Freedom. 3. Integration of Equations of Small Oscillations. 4. Review of Operational Calculus. 5. Forced Oscillations of a System with a Finite Number of Degrees of Freedom. 6. Approximate Methods for Determining the Fundamental Frequency. 7. Approximate Methods of Determining the Higher Frequencies. 8. Resonance Phenomena in Machinery.

Part 2. Linear Systems with Infinite Number of Degrees of Freedom. Chapters: 9. General Properties of Small Oscillations of Elastic Rods. 10. Longitudinal and Torsional Oscillations of Straight Rods. 11. Transverse Oscillations of Straight Rods. 12. Approximate Methods of Calculation of Oscillations in Straight Rods of Variable Cross Section. 13. Transverse Oscillations of Plates.

Part 3. Stability of Motion and Nonlinear Oscillations. Chapters: 14. Introduction to the General Theory of Stability of Motion. 15. Stability in the First Approximation. 16. Simplest Nonlinear Systems. 17. Certain General Methods of Nonlinear Mechanics. Appendix: Tables, Author Index, Subject Index.

Beck, A. Thermionic Valves, Their Theory and Design. Translated from the English by K. N. Trofimov and I. E. Ovsievich, edited by L. A. Kotomina, Moscow, "Soviet Radio" Press, 1958. [Cambridge University Press, 1953.]

Blokhin, M. A. Физика рентгеновских лучей. (Physics of X-rays) Second edition, revised, Moscow, Gostekhizdat, 1957, 518 pp. illustr., bibliography at the end of the book by chapters (347 titles). 7,000 copies, 17.50 rub.

Bruining, H. Physics and Application of Secondary Electron Emission. Translated from the English by B. S. Kul' varskaya and G. V. Stepanov, edited by V. N. Favorin, Moscow, "Soviet Radio" Press, 1958, 8.05 rub. [McGraw-Hill, New York, 1954.]

Vainshtok, I. S. Ультразвук и его применение в машиностроении (Ultrasound and Its Application to Machine Building) Moscow, Mashgiz, 1958, 140 pp., illustr., one set of diagrams. Bibliography pp. 134-138 (109 titles). 8,000 copies, 4.65 rub.

*Books marked with an asterisk were placed on sale in 1958.

Chapter headings: 1. Physical Principles of Commercial Application of Ultrasound. 2. Methods of Generating Ultrasonic Oscillations. 3. Phenomena Arising in Substances under the Influence of Ultrasound. 4. Ultrasonic Apparatus. 5. Ultrasonic Mechanical Finishing. 6. Ultrasonic Soldering and Tinning. 7. Ultrasonic Cleaning of Parts. 8. Possible Applications of Ultrasonics in Metallurgy. 9. Protection of Boilers Against Scale Formation and Sonic Coagulation of Aerosols. 10. Ultrasonic Defectoscopy. 11. Ultrasound in Material Research. 12. Ultrasonic Measurements.

Ventsel', E. S. Теория вероятностей (Theory of Probability) Moscow, State Publishing House for Physical and Mathematical Literature, 1958, 464 pp, diagrams, one sheet of tables. 15,000 copies, 9.55 rub. in case.

Chapter headings: 1. Introduction. 2. Fundamental Concepts of the Theory of Probability. 3. Principal Theorems of the Theory of Probability. 4. Repetition of Experiments. 5. Random Quantities and their Distribution. 6. Normal Distribution Law. 7. Determination of the Distribution Law of Random Quantities on the Basis of Experimental Data. 8. Systems of Random Quantities. 9. Normal Law of Distribution for a System of Random Quantities. 10. Numerical Characteristics of Functions of Random Quantities. 11. Linearization of Functions. 12. Laws of Distribution of Functions of Random Arguments. 13. Limit Theorems in the Theory of Probability. 14. Processing of Experiments. 15. Fundamental Concepts of the Theory of Random Functions. 16. Canonical Expansions of Random Functions. 17. Stationary Random Functions. Appendix 1. Tables. 2. Grid of Normal Scattering Law. Subject index.

Вопросы электролюминесценции (Problems of Electroluminescence) (Translations of articles). Moscow-Leningrad, Gosenergoizdat, 1958, 32 pages with illustrations (Foreign Electronics). Bibliography at the end of each article. 7100 copies, 1.60 rub.

Voronchev, T. A. Импульсные тиратроны (Pulsed Thyratrons), Moscow, "Soviet Radio" Press, 1958, 164 pp, diagrams. Bibliography pp 161-162 (36 titles), 6.35 rub.

Chapter headings: 1. Internal Information on Pulsed Thyratrons. 2. Physical Phenomena in Pulsed Thyratrons During the Part of the Cycle Prior to Discharge. 3. Physical Phenomena in Pulsed Thyratrons During the Conducting Portion of the Cycle. 4. Physical Phenomena in Pulsed Thyratrons in the Post-Discharge Portion of the Cycle. 5. Analysis of the Circuits of a Linear Modulator with DC Supply. 6. Technological and Structural Features of Pulsed Thyratrons.

*Всесоюзное совещание по спектроскопии (Tenth All-Union Conference on Spectroscopy) L'vov, 1956. Transactions (editorial board headed by G. S. Landsberg) L'vov University Press, 1957 (Ministry of Higher Education, Ukrainian S.S.R., the I. Franko L'vov State University, Physics collection...). Vol. 1. Molecular Spectroscopy, 409 pp. illustr. 17 inserted illustrations [No. 3(8)]. Bibliography at the end of each article, 4,000 copies, 33 rub.

*Дозиметрическая, радиометрическая и электронно-физическая аппаратура (Dosimetric, Radiometric, and Electron-Physical Apparatus) (Editor in chief, I. G. Kokin), Moscow, Atomizdat, 1957, 99 pp. illustr. (Ministry of Radiotechnical Industry, U.S.S.R., Scientific Research Institute, Information Collection No. 1). 2,000 copies, no price.

Contents: 1. Electron-Physical Instruments for Nuclear Research. 2. Sanitary-Dosimetric Instruments. 3. Instruments for Geological Prospecting. 4. Other Instruments.

Zaborenko, K. B. Радиоактивность (Radioactivity). Edited by Distinguished Scientist Professor V. I. Baranov. Second, revised edition. Moscow, Gostekhizdat, 1958, 80 pp. illustr. (Scientific Popular Library, No. 54). List of publications in the series "Scientific Popular Library, 1945-1957" on pp 78-80. 75,000 copies, 1.20 rub.

Shielding of Nuclear Reactors. Translated from the English and edited by S. G. Tsypin, Moscow, Foreign Literature Publishing House, 1958, 344 pp. illustr. two sheets of graphs (Reports of the Atomic Energy Commission, U.S.A.). Bibliography at the end of each chapter. 26.40 rub.

Van der Ziel, A. Noise in Radio Engineering and Physics. Translated from the English, edited by L. S. Gutkin. Moscow-Leningrad, Gosenergoizdat, 1958.

Исследования по люминесценции (Research on Luminescence) Collection of articles, editorial board headed by F. D. Klement. Tartu, 1958, 362 pp., graphs (Academy of Sciences, Estonian S.S.R. Transactions of the Institute of Physics and Astronomy, No. 7). Summary of articles in English, bibliography at the end of each article. 800 copies.

*Kaufman, M. S., and Yankin, G. M. Электронные приборы (Electron Tubes). Edited by Professor R. A. Nilender. Part 2, Moscow-Leningrad, Gosenergoizdat, 1957, 320 pp. illustr. 15,000 copies, 7.40 rub.

Chapter headings: 10. Receiving and Amplifying Tubes. 11. Oscillator Tubes. 12. Electron Tubes for Microwaves. 13. Klystrons, Magnetrons, and Traveling and Backward Wave Tubes. 14. Photoelectronic Tubes. 15. Cathode-Ray and Electron-Optical Tubes. 16. X-ray Tubes. 17. Semiconductor Rectifiers and Amplifiers. Dischargeless Devices.

Kitaigorodskii, A. I. Теория структурного анализа (Theory of Structural Analysis), Moscow, U.S.S.R. Acad. Sci. Press, 1957, 284 pp., diagrams. 5,000 copies, 10.65 rub.

Chapter headings: 1. Mathematical Introduction. 2. Fundamentals of the Theory. 3. Structural Amplitudes and Products as Random Quantities. 4. Theory of Connection of Structural Amplitudes. 5. Investigation of Contraction of Electron Density. 6. Methods Correlating the Measured and Calculated Structural Amplitudes. Conclusions.

Komarovskii, A. N. Защитные оболочки ядерных реакторов (Protective Shields for Nuclear Reactors) Moscow, Atomizdat, 1958, 67 pp. illustr., bibliography pp 66-67 (49 titles). 5,650 copies, 2.70 rub.

Kothari, Homi-Bhabha, and Khanlokar, Nuclear Explosions. Translated from the English by N. F. Kravtsova, edited by N. P. Dushnov, Moscow, Foreign Literature Press, 1958.

Short German-Russian Dictionary on Nuclear Physics and Nuclear Engineering. Compiled by Yu. M. Kaplanskaya, A. M. Lidvanskii, and N. F. Manushin, edited by Doctor of Technical Sciences D. I. Voskoboinik. Moscow, Gostekhizdat, 1958, 303 pp., 20,000 copies, 8.90 rub. (Compiled from materials of the Institute of Scientific Information of the Academy of Sciences, U.S.S.R.).

Krize, S. N. Усилительные устройства (Amplifier Circuits) Textbook for Communication High Schools. Moscow, Svyaz'izdat, 1958, 315 pp. illustr. Bibliography p. 304 (16 titles). 50,000 copies, 7.40 rub.

Contents: Part 1. General Information on Amplifier Circuits. 2. Power Amplifiers. 3. Voltage Amplifiers.

Krichevskii, E. S., Fedorovich, L. G., and Fetisov, V. F. Электрооборудование оптико-механических приборов (Electric Equipment for Optical-Mechanical Instruments) Textbook for Technical High Schools. Moscow, Oborongiz, 1958, 468 pp. illustr. seven sheets of diagrams. Bibliography pp. 460-464. 8,000 copies, 12.20 rub.

Chapter headings: 1. Electrotechnical Materials Used for the Manufacture of Parts and Components of Electric Equipment for Instruments. 2. Electric and Electronic Parts Used in Electric Equipment for Instruments. 3. Electric Motors and Sources of Supply Used in Electric Equipment for Instruments. 4. Electron Tubes and Amplifier Circuits. 5. Current and Voltage Stabilizers. 6. Radiation Sources and Receivers. 7. Electric Heating in Optical-Mechanical Instruments. 8. Principles of Electron Optics and its Technical Application. 9. Electric Diagrams of Instruments.

Krugman, L. M. Transistors and Their Applications. Translated from the English by M. A. Barg. Moscow-Leningrad, Gosenergoizdat, 1957. [N.Y., Rider, 1955].

Kuznetsov, B. G. Принципы классической физики (Principles of Classical Physics) Moscow, U.S.S.R. Acad. Sci. Press 1958, 323 pp. (Academy of Sciences, U.S.S.R., Institute of History of Natural Sciences and Technology). Bibliography and footnotes. 7,000 copies, 13.20 rub.

Contents: 1. Introduction. 2. The Principle of Relativity. 3. The Principle of Least Action. 4. The Principle of Conservation of Energy. 5. The Principle of Irreversibility. 6. The Principle of Near Action.

Lazarev, P. P. Сочинения (Written Works) (Editorial commission headed by S. I. Vavilov. Introductory article by V. V. Shuleykin. Biographical outline by B. V. Deryagin and M. P. Volarovich). Vol. 1, Moscow-Leningrad, Published U.S.S.R. Acad. of Sci. Press 1957.

Vol. 1. (Works on biophysics. Research in adaptation. Outline of the history of Russian science. Edited and introductory article by B. V. Deryagin). 895 pages with illustrations, nine sheets of illustrations. "Bibliography of the works of P. P. Lazarev, pp. 855-891; "Annotation of the Works Not Contained in the Present Publication" pp 823-853, compiled by M. P. Volarovich, V. V. Efimov, B. V. Deryagin, and S. S. Kovner. Bibliography, pp. 39-40 (25 titles), 35,000 copies, 42.90 rub.

Landsberg, Grigorii Samuilovich, Избранные труды (Collected Works) Edited by Doctor of Physical-Mathematical Sciences I. L. Fabelinskii (Introductory article by S. L. Mandel'shtam and I. L. Fabelinskii, pp. 5-40). Moscow, U.S.S.R. Acad. Sci. Press 1958, 476 pp. illustr., one por-

trait. 4,000 copies, 27.45 rub. Bibliography of the works by G. S. Landsberg and relevant literature, pp. 464-474, compiled by T. O. Vreden-Kobetskoi. Bibliography at the end of each chapter.

Contents: 1. Molecular Scattering of Light. 2. Raman Scattering of Light. 3. Applied Spectroscopy (Atomic and Molecular). 4. Various Papers. 5. Survey Papers and Unpublished Manuscripts.

*Landsberg, G. S., Оптика (Optics), Textbook for State Universities. Fourth revised edition. Moscow, Gostekhizdat, 1957, 759 pp. illustr. (Physics General Course, Vol. 3), 50,000 copies, 15.65 rub.

L'vov, V. E., Жизнь Альберта Эйнштейна (Life of Albert Einstein) 1879-1955. Moscow, Molodaya Gvardiya (Young Guard) 1958, 320 pp. illustr. seven sheets of illustrations (Lives of Outstanding Persons). Bibliography pp. 316-319. 15,000 copies, 6.55 rub.

*Lyubarskii, G. Ya., Теория групп и ее применение в физике (Group Theory and Its Application in Physics) Moscow, Gostekhizdat, 1959, 354 pp. 6,000 copies, 13.20 rub.

Chapter headings: 1. Elements of Group Theory. 2. Certain Specific Groups. 3. Theory of Representation of Groups. 4. Operations with Commutation of Groups. 5. Commutation of Certain Groups. 6. Small Oscillations of Symmetric Systems. 7. Second-Order Phase Transitions. 8. Crystals. 9. Infinite Groups. 10. Representations of the Rotation Groups and of the Complete Orthogonal Group. 11. The Clebsch-Gordan and Racah Coefficients. 12. The Schrödinger Equation. 13. Equations Invariant with Respect to the Euclidian Group of Space Motions. 14. Absorption and Raman Scattering of Light. 15. Representations of the Lorentz Group. 16. Relativistically Invariant Equations. 17. Nuclear Reactions.

Appendices: Bibliography at the end of each chapter, subject index.

Migulin, V. V., Лекции по основам радиолокации (Lectures on the Principles of Radar) Moscow, Moscow University Press, 1958, 123 pp., diagrams, 26,000 copies, 3.10 rub.

Contents: Foreword. Introduction. Chapters: 1. General Principles of Operation of Radar Systems. 2. Basic Energy Relations. 3. Influence of Radio-Wave Propagation Conditions. 4. Operation of Antennas in Transmission and Reception. 5. Reflection and Scattering. 6. Noise in Receivers and Maximum Sensitivity of Receivers. 7. Principle Methods of Determining the Direction of the Reflecting Object (Direction Finding). 8. Radar Measurements of Range Using Frequency Modulation. 9. Measurements of Range Using Pulse Modulation and the Operating Features of Pulse Systems. 10. General Characteristics of Radar Systems. 11. Selection of Moving Targets.

International Conference on Peaceful Use of Atomic Energy, Geneva, 1955. Transactions of the International Conference on Peaceful Use of Atomic Energy, held in Geneva 8 to 20 August 1955 (United Nations). Vol. 5. Physics of Reactors. 1958. Moscow, U.S.S.R. Acad. Sci. Press, 1958.

Bethe-Hoffman. Mesons and Fields (in two volumes). Translation from the English edited by Academician I. E. Tamm, Vols. 1-2, Moscow, Foreign Literature Press, 1957.

Massey, G. and Barhop, E., Electron and Ion Collisions. Translation from the English edited by S. M. Osovets. Moscow, Foreign Literature Publishing House, 1958. The book contains also "Theory of Scattering of Slow Electrons" by G. Massey.

Mitsusima, Sanitiro, Structure of Molecules and Internal Rotation. Translation from the English by V. M. Gryaznov and V. D. Yagodovskiy. Edited by Professor V. M. Tatevskiy. Moscow, Foreign Literature Publishing House, 1957.

Danger to Humans from Ionizing Radiation. Translated from the English by A. E. Kal'manson. Foreword to the Russian edition by Professor F. G. Krotkov. Foreign Literature Publishing House, 1958.

The book is intended for doctors, biologists, physicists, engineering-technical workers, and all persons dealing with sources of ionizing radiation.

*Orestov, I. L., Холодный свет (Люминесценция) [Cold Light (Luminescence)]. Second edition, Moscow, Gostekhizdat, 1957, 40 pages with illustrations (Scientific-Popular Library, No. 82) 50,000 copies, price 0.60 rubles.

*Fundamental Formulas of Physics. Edited by D. Menzel. Translated from the English under the editorship of I. S. Shapiro. Moscow, Foreign Literature Press, 1957. [Prentice Hall, 1955.]

Fundamentals of Transistor Electronics. Translated from the English under the editorship of E. I. Gal'perin. Moscow, "Soviet Radio" Press, 1958. [No reference to author or editor of original].

Petrov, I. N., Полупроводниковые приборы (Transistor Devices). Moscow, Voyenizdat, 1957, 128 pp., bibliography. 2.25 rubles.

Introduction. 1. Fundamental Properties of

Semiconductors. 2. Types and Constructions of Semiconductor Devices. (The purpose of this brochure is to explain the operating principle and the construction of several semiconductor devices, and describe the physical processes that occur in semiconductors).

Petrov, P. A., Ядерные энергетические установки (Nuclear Power Installations). Moscow-Leningrad, Gosenergoizdat, 1958, 526 pp. illustr. Bibliography at the end of each chapter. 15,000 copies, 10.75 rubles.

Chapter headings: 1. Basic Concepts of Nuclear Physics. 2. Nuclear Reactions. 3. Physics of Nuclear Reactors. 4. Operation of Nuclear Reactors. 5. Shielding Against Radiation. 6. Removal of Heat From Nuclear Reactors. 7. Materials of the Active Zone of the Reactor. 8. Construction and Arrangements of Nuclear Power Installations. 9. Use of Nuclear Reactors for the Generation of Electricity. Appendices: 1. Calculation of a Nuclear Power Reactor. 2. Effective Absorption Cross Sections for Thermal Neutrons. 3. Nomogram for Determination of Quantities. 4. Prices of Nuclear Materials in the U.S.A.

Получение изотопов. Мощные гамма-установки. Радиометрия и дозиметрия. (Production of Isotopes. High Power Camera Ray Installations. Radiometry and Dosimetry). Collection of Articles. Editorial staff: Yu. S. Frolov (Editor in chief) and others. Moscow, U.S.S.R. Acad. Sci. Press, 1958, 294 pp., illustr. (Academy of Sciences, U.S.S.R., Main Administration for the Use of Atomic Energy at the Council of Ministers U.S.S.R. Transactions of the All-Union Scientific-Technical Conference on the Use of Radioactive and Stable Isotopes and Radiations in the National Economy and in Science, 4-12 April, 1957).

The binding has the following heading: All-Union Conference on Use of Isotopes and Nuclear Radiation. Bibliography at the end of each chapter, 5,000 copies, 17.35 rub.

***Pohl, R. W.** Mechanics, Acoustics and the Theory of Heat. Translation from the 13th German edition by K. A. Leont'ev and V. M. Yuzhakov, rev. and suppl. Edited by N.P. Suvorov. Moscow, Gostekhizdat, 1957. [Springer, 1955].

Wright, D. Semiconductors. Translated from the English by B. Ya. Moizhes, edited by S. S. Shalyt. Moscow, Foreign Literature Press, 1957. [Methuen, London, 1955.]

***Russell, Bertrand.** Human Knowledge, Its Scope and Limits. Translated by N. V. Vorob'yev. Editor and author of introductory article, E. Kol'man. Moscow, Foreign Literature Press, 1957.

Rizkin, A. A. Основы теории усилительных схем (Fundamentals of the Theory of Amplifier Circuits). Third edition, revised and supplemented. Moscow, "Soviet Radio" Press, 1958, 527 pp. illustr. 18.80 rub.

Chapter headings: 1. General Methods of Analysis of Amplifier Circuits. 2. Selective Amplifiers. 3. Voltage Amplifers for Audio Frequencies. 4. Feedback Amplifiers. 5. Certain Feedback-Amplifier Circuits. 6. Broadband Amplifiers. 7. Pulse Amplifiers. 8. Audio Power Amplifiers. 9. Transistor Amplifiers. 10. Special Types of Amplifiers.

Sutton, O. G. Micrometeorology. Investigation of Physical Process in the Lower Layers of the Atmosphere. Translated from the English under the editorship of D. L. Laykhtman. Leningrad, Gidrometeoizdat, 1958.

Slyusarev, G. G. О возможном и невозможном в оптике (The Possible and Impossible in Optics). Second edition, revised. Moscow, Gostekhizdat, 1957, 178 pp. illustr. 12,000 copies, 2.90 rub.

Spitzer, L. Physics of Fully Ionized Gases. Translated from the English under the editorship of R. A. Demirkhanov. Moscow Foreign Literature Press, 1957.

*Telemetering and Remote Control. Transactions of the National Conference on Telemetering held in Chicago 1954. Collection of articles, translated from the English under the editorship of S. V. Alekseev and A. V. Lebedev. Foreign Literature Press, 1957.

Философские вопросы современной физики (Philosophical Problems of Modern Physics) Collection of articles and translations. Edited by I. V. Kuznetsov and M. E. Omel'yanovskii. Moscow, Gospolitizdat, 1958, 248 pp. 15,000 copies, 7 rub.

Contents: M. E. Omel'yanovskii, Dialectical Materialism and Modern Physics. B. M. Kedrov, On the Classification of Sciences. V. A. Fock, Critics of Bohr's Views on Quantum Mechanics. Louis De Broglie (France), Interpretation of Wave Mechanics. V. S. Sorokin, Law of Conservation of Motion and Measure of Motion in Physics. Thomas A. Brody (Mexico) Formation and Region of Applicability of Scientific Concepts. N. V. Markov, Philosophical Significance of the Theoretical Inheritance of N. I. Lobachevskii. Wladyslaw Krajewski (Poland), The Struggle of Marjan Smoluchowski for Scientific Atomistics. E. Kol'man, What is Cybernetics?

Fogel'son, B. A. Волноводы (Waveguides). Moscow, Voyenizdat, 1958, 125 pp. illustr. (Radar Engineering), 2 rub.

Contents: I. Waves and Wave Motion. 1. General Information. 2. Phase and Group Velocities of Waves. 3. Electromagnetic Waves. II. Guiding Properties of Metallic Plane. 2. Waves Between Parallel Planes. III. Hollow Waveguides. 1. Rectangular Waveguides. 2. Round Waveguides. IV. Elements of Waveguide Systems. 1. Stationary Junctions. 2. Movable Junctions. 3. Matching Elements. 4. Branches. V. Connection Between Waveguide Systems and HF Sources and Receivers. Appendix.

Frenkel', Ya. I. Введение в теорию металлов (Introduction to the Theory of Metals) edited by S. V. Vonsovskii. Third edition, Moscow, State Publishing House for Physical-Mathematical Literature, 1958. 368 pp., illustr. ("Physical-Mathematical Engineering Library") Bibliography at the end of each chapter, 8,500 copies, 11.60 rub.

Frish, Ē. S. and Timoreva, A. V. Курс общей физики (Textbook of General Physics) Text for universities. Moscow, State Publishing House for Physical-Mathematical Literature, 1958, Vol. 1. Physical Fundamentals of Mechanics. Molecular Physics. Vibrations and Waves. Ninth Stereotyped Edition, 1958, 463 pp., illustr. 100,000 copies, 10.60 rub.

Kharkevich, A. A. Спектры и анализ (Spectra and Analysis). Third revised edition, Moscow, Gostekhizdat, 1957, 236 pp. Bibliography, 32 titles. 10,000 copies, 6.80 rub.

Contents: 1. Spectra. 2. Analysis. 3. Spectra of Random Processes. Appendices (Many additions and corrections have been made in the third edition).

Sharonov, V. V. Природа планет (The Nature of the Planets). Moscow, State Publishing House for Physical-Mathematical Literature. 1958, 552 pp., illustr. and maps. Bibliography at the end of each chapter. 30,000 copies, 19.05 rub.

Chapter headings: 1. Introduction. 2. Telescopic Investigation of the Surface of Planets and Satellites. 3. Geometrical and Mechanical Characteristics of the Solar System. 4. Topography and Cartography. 5. Topographic Description of Individual Members of the Solar System. 6. Integral Photometer. 7. Photometry of the Planet Disks. 8. Optics of Planetary Atmospheres. 9. Physical Conditions on the Planets and Satellites. Alphabetical Index.

Schulz, M. A. Control of Nuclear Power Reactors. Translated from the English by I. B. Vikhanskii and V. V. Korolev, edited by D. I. Voskoboinik. Moscow, Foreign Literature Publishing House, 1957. [McGraw-Hill].

Shuppe, G. N. Электронная эмиссия металлических кристаллов : (Electron Emission of Metallic Crystals) Tashkent, Press of the Central-Asia University, 1957, 112 pp., illus. (Ministry of Higher Education, U.S.S.R., V. I. Lenin Central-Asia State University, Transactions, New series, No. 115, Physical-Mathematical Sciences, Book 17). Bibliography pp. 109-111 (141 titles). 500 copies, 5-rub.

Элементарный учебник физики (Elementary Textbook of Physics) edited by Academician G. S. Landsberg. Second edition, revised. Vols. 1-3, Moscow, State Publishing House of Physical-Mathematical Literature. 1958, Vol. 1. Mechanics. Heat. Molecular Physics, 1958, 523 pp., illustr. 50,000 copies (reprint), 12 rub. Vol. 2. Electricity and Magnetism. 1958, 448 pp. illustr. 50,000 copies (reprint), 9.95 rub.

Andrew, E. Nuclear Magnetic Resonance. Translation from English by N. M. Pomerantsev and E. N. Skubur. Edited by V. N. Lazukin. Moscow, 1957. [Cambridge, 1955].

Yakobi, B. S. Работы по электрохимии (Papers on Electrochemistry). Collection of articles and materials edited by A. N. Frumkin. Introductory article by editor and comments by M. I. Radovskii. Moscow-Leningrad. U.S.S.R. Acad. Sci. Press. 1957, 304 pp (Academy of Sciences, U.S.S.R. Institute of History of Natural Sciences and Technology). Bibliography in the notes on pp. 192-207. 3,000 copies, 13.15 rub.

Contents: A. N. Frumkin, B. S. Yakobi's Work in the Field of Electrochemistry. Texts in the appendices: I. G. Spasskii, The First Years of Galvanoplastics in Russia. Materials on the History of the Development and Introduction of Galvanoplastics.

— T. O. Vreden-Kobetskaya

Usp. Fiz. Nauk **66**, 148-156 (September, 1958)

G. Hertz (Herausgegeben von). Lehrbuch der Kernphysik. Band I. Experimentelle Verfahren, pp 227. B. G. Teubner Verlagsgesellschaft, Leipzig, 1958.

The book reviewed is the first volume of an interestingly conceived collective* textbook of nuclear physics, which when completed will consist of three volumes. The second volume will be devoted to the physics of the nucleus, and the authors plan for it to play the central role. It is assumed that along with the development of the factual material, a proper place will also be allotted to theoretical problems. Finally, the third volume will be devoted to applied nuclear physics: reactors, isotope separation, radiochemistry, and the use of stable and radioactive isotopes. In the editor's foreword Professor G. Hertz states the purpose of the entire undertaking in the following manner. Modern nuclear technology is progressing rapidly, but the development of nuclear physics is far from complete. Under these conditions it becomes very important to train a large number of physicists as specialists in the field of nuclear physics. It is for physicists specializing in this field that this book is intended. The book can also be useful to engineers educated in other fields of engineering and who are retraining to become nuclear engineers. This textbook is thus intended for all these readers, and also for physicists who are working in other fields but who make use of the methods of nuclear physics for special purposes. The authors had in mind to write not an elementary introduction and not a collection of special monographs, but a serious textbook for a basic treatment of the subject. The task so formulated is exceedingly difficult. The editor, the outstanding physicist G. Hertz, who as the partner in the famous "Frank and Hertz experiment" is now known even to school boys interested in physics, has naturally understood it only too well and one must therefore agree with his most correct opinion, expressed in the foreword, that each part of such a textbook should be written by a specialist engaged in this problem or one who is developing the particular experimental procedure.

A specialist can readily, and with suitable emphasis, choose the most important fresh material and treat it didactically in the most suitable manner.

The content of the reviewed first volume is as follows. An introduction of 36 pages describes very briefly the modern ideas concerning the atomic structure and the principles of quantum theory. The remainder of the book is divided into three parts. Part A (pp 37-103) is devoted to methods of experimental determination of the principal characteristics of the nucleus; charge, radius, mass, and momenta. The methods considered for the determination of the radius of the nucleus are those connected with the action of nuclear forces (α-decay periods, effective cross section for the scattering of α particles and fast neutrons, etc.), as well as those connected with the electrostatic interactions (β-decay energy, scattering of fast electrons, μ-mesic atoms). The theoretical problems are really not considered, and the formulas are given without derivations, with reference (in some cases) to the second volume. The next to be considered are methods of determining the atomic mass. In this part (pp 48-79) there is a very well grounded survey of modern corpuscular optics and mass spectrometry. Again, derivations are given here only for the simplest cases. To master this section completely the reader will probably have to resort to more extensive books and monographs devoted to corpuscular optics.

The last part of Part A is devoted to determination of the nuclear moments. All modern methods are considered: magnetic and quadrupole resonance, hyperfine structure of atomic spectra, molecular spectra, and atomic beams. The theoretical Sect. 11 is devoted to the interaction between the nucleus and electric and magnetic fields, a topic important to the understanding of the methods later described. The concept of electric and magnetic multipoles and of dipole and quadrupole interactions is introduced here. As is known, these are cumbersome problems, difficult to master. However, the treatment is so compact that it will probably be inaccessible to the average student.

Part B is devoted to methods of observation of nuclear particles. The usual methods of observation and detection of fast particles are considered: all types of counters, cloud chamber, and nuclear emulsions. Modern improvements of cloud chambers — diffusion and bubble chambers — are barely mentioned. Various electrometers used for the measurement of ionization current are described and the principles of radio electronics are expounded. The last paragraph of this part is devoted to the spectrometry of α and β particles.

Part C, the last, is devoted to elementary-particle accelerators. Compared with the preceding parts of the book, the exposition is more descriptive and therefore also more elementary.

The authors mention but few names in the book

*The reviewed volume was compiled by Dr. F. Bernard, Prof. A. Eckard, Prof. W. Harman, Prof. G. Hertz, A. Loesche, Prof. I. Schintlmeister, and Prof. K. Weiss.

and in general this is proper in a textbook. The Soviet reader will notice with satisfaction due credit to papers and names of Soviet physicists, with one very surprising exception. As is known, the principle of self-phasing, which plays a decisive role in the construction of modern resonant accelerator, was first formulated by the Soviet physicist V. I. Veksler and later, independently by the American physicist MacMillan. This situation, particularly the priority of V. I. Veksler, has long been acknowledged in the entire world, in the literature of almost all countries. Yet the compiler of the chapter on accelerators (Professor A. Eckardt), while readily mentioning authors of much less important papers, limits himself to an indirect reference to Veksler in connection with a fact that is rather special and of no principal significance (nor is, incidentally, MacMillan mentioned). The purpose of this clearly intentional silence on work that has influenced decisively the modern techniques of obtaining particles with super high energies is quite puzzling.

As a whole the book can be classified as a useful summary. The two bibliographies at the end of the book, a list of original papers and a list of scientific books and monographs, make it easier for the reader, who will undoubtedly have to resort to simultaneous reading of other books and articles.

— E. V. Shpol'skii

D. A. Wright. Semiconductors. Translated from the English by B. Ya. Moizhes, edited by S. S. Shalyt. Foreign Literature Press, 1957. [Methuen, 1955].

Modest in size (154 pages of small format), Wright's book contains a very readable exposition of many problems in semiconductor physics. The first two chapters contain an outline of the principal concepts of electron theory of metals and semiconductors (within the framework of the usual Sommerfeld and Bloch models, with warning concerning the difficulties of the band theory as applied to substances of the nickel-oxide type). The exposition is a successful one. The author gives a clear picture of the phenomena, without resorting thereby to the complex mathematical formalism of modern theoretical physics. Naturally, the principal formulas, which must nevertheless be used, can be introduced in such a treatment only purely dogmatically; however, the availability of many good textbooks (to which reference is made) makes this quite permissible in a book of this type.

Next to be considered is the emission of electrons from the surface of a solid, the determination of the concentration of electrons on semiconductors, secondary emission, contact between a metal and a semiconductor, thermionic cathodes, and photocathodes. This list of chapters already shows that the book is written in a somewhat different plan than has become customary in recent times. The principal emphasis is on cathode electronics and allied problems, and here the author develops his argument in considerable detail (although, apparently, somewhat subjectively), sometimes perhaps even overloading it with details. At the same time problems of principal and technical importance to the physics of contact phenomena, as well as kinetic problems, are discussed at best sketchily; some important problems (such as recombination of carriers in a semiconductor) are not considered at all. Apparently this is due to the author's personal interests. To some extent such subjectivity of exposition is naturally unavoidable; nevertheless, in this case the emphasis on cathode electronics (for all its importance) must be considered excessive. This is more so regrettable, since the book demonstrates splendidly the author's ability to expound briefly and clearly even comparatively difficult topics. Were the material more uniformly distributed, Wright's work could serve as a beautiful introduction to modern semiconductor physics. However, even in its present form it can be useful to engineers interested in elements of this science. On the other hand, the "cathode" portion of the book, which illuminates an important and interesting region to which little attention has undeservedly been paid heretofore in our popular literature, will interest also specialists in solid state physics.

As a whole, Wright's book must be favorably reviewed and its translation into Russian is fully justified. No special complaints can be made concerning the quality of the translation.

— V. Bonch-Bruevich

Usp. Fiz. Nauk **66**, 349-351 (October, 1958)